Université de Bourgogne
Laboratoire d'Etude de l'Apprentissage et du Développement
C.N.R.S. – U.M.R 5022

Ecole Doctorale Environnement, Santé, STIC (E2S)

Comprehension of Complex Animation: Cueing, Segmentation, and 2D/3D presentations

Thèse de Doctorat de l'Université de Bourgogne

- Mention Psychologie –

Présentée par **Dian Kemala PUTRI**

Sous la direction du Professeur **Jean-Michel BOUCHEIX**

<u>Membres du Jury :</u>

Professeur BASTIEN Jean-Christian (Rapporteur -Université Paul Verlaine, Metz-France)

Professeur BETRANCOURT Mireille (Rapporteur -Université de Genève-Suisse)

Professeur BONIN Patrick (Président du jury-Université de Bourgogne-France)

Professeur BOUCHEIX Jean-Michel (Directeur de Thèse, Université de Bourgogne-France)

Professeur MERIENNE Frédéric (Membre du jury, ENSAM-Institut Image, Chalon-sur-Saône, France)

Professeur FOULIN Jean Noel (Membre du jury, Université de Bordeaux 4 (IUFM), Bordeaux, France)

- 25 Novembre 2011 -

To my dearest husband: Tanto Pratondo Utomo
my wonderful children:
Dito and Laras

For their love, ever patient and endless support ...

ACKNOWLEDGEMENTS

I am heartily thankful to the supervisor of my dissertation, Professor Jean Michel Boucheix whose encouragement, guidance, and support from the initial to the final level and enabled me to develop an understanding of the subject. Thank you for your kindness and your patience to introduce me with the ‘new world’ which I never realized before...Thank you for the belief on my work.

A great appreciation to the members of my committee and jury, Professor Frédéric Merienne for giving me the opportunity to pursue doctoral program in France and introduced me to Professor Jean Michel Boucheix. I would like to thank for Professor Mireille Betrancourt and Professor Jean-Christian Bastien who accepted to review my work and give invaluable input and also to Professor Patrick Bonin as president of jury and Professor Frédéric Merienne and Professor Jean Noel Foulin as member of jury.

A great appreciation to Professor Dr.E.S.Margianti,S.E.,MM as my rector at Gunadarma University for her great support dealing with my Ph.D Program.

I would like to thank for the Indonesian Ministry of Education, Directorate General of Higher Education who supported with the scholarship and also L’Ecole Doctorale de L’Université de Bourgogne and the LEAD for the “bourse de mobilité” when I participated in international conferences during my PhD years (APCHIERGO2010-Bali and EARLI2011 Exeter-UK).

I acknowledge to Juan Sebastian Casallas for his OCNOSCERE software, you are a great partner in my research with eye tracking devices, thanks for discussion, ideas and it is pleasant having a great friend like you. I also wish to acknowledge to Blandine Brunel, you are also my best team work in research and Stéphane Argon, for helping to create software supported in Tobii for our experiments.

I would like to thank the Secretaries of LEAD (Sandrine Pinson and Corinne Martin) for their help, their encouragement; they always support me to finish my work. Thank you also to Elodie, Lana, Milena, Gérome, Diana and all my friends for companionship and colleagues in LEAD.

I am grateful to Dr. Boone Sumantri, my uncle who helped me for clarifying my English grammar and enriched my vocabulary in writing the thesis.

Thank you to Ira, Cut, Aries, Musa, Okta, Iqbal, Aning, and Fitrin, for all your caring ways, and all my Indonesian friends thanks for always support me and keep spirit to do the best thing and many thankful to Franco-Indonesian families in Dijon.

Lastly, I offer my regards and blessings to my great family, beloved father (in memoriam), mom, mother in law, sisters and brothers who understand and support me in any respect during my PhD. My dearest husband Tanto Pratondo utomo for your trust, love, patience and always support me and available although separated by long distance…My wonderful children Dito Satrio Utomo and Adinda Larasati Utomo…you are my big engine, giving me spirit to keep on and always say: "*never give up mom*"...really you are the best I have… I want to thank Allah SWT who made all things possible and for providing for me in every way...

Abstract

The goal of our studies was to test the effect of segmentation, cueing, and 2D/3D presentations to foster complex animation processing. The material was an upright mechanical piano system. We used an eye tracking system which provides information about learners' attention direction during the animation processing. We analyzed the effect of the format presentations and the eye movements during learning.

Based on animation and multimedia research background, four experiments were conducted. In the first experiment the effect of the presentation of simplified external representations on learning from complex animation was investigated. Experiment two and three aimed at studying the cognitive processes involved in learning to complex mechanism system with new cueing techniques with spatial-temporal colored tokens. In the fourth experiment, 2D and 3D presentation of the same animated content were compared.

Results of these experiments showed that (1) the use of a dual format presentation is better for developing a dynamic mental model from the animation than a single format, (2) the signaling strategies using cued tokens of dual format can guide efficiently learner's building of mental model and can enhance learner's comprehension of complex system, (3) a sequential format presentation followed by an animation format presentation helps the learner to understand the key stages of a dynamic process and to create a high quality mental model, (4) 3D animation presentation is better than 2D animation presentation to direct attention on relevant component of the animation. For depth processing, comprehension with 3D animation presentation is better than 2D animation format. Eye tracking measures provided insights into how people understood and learned complex systems. The analysis of eye tracking also contributed to the understanding of the subject's perceptual processing during learning.
Overall, results provide a significant contribution in the field of learning with complex animation. Recommendations in the ergonomics area for the design of animations are proposed.

Keywords: animation, segmentation, cueing, tokens, 2D/3D presentations, eye tracking

Table of contents

List of tables

List of Figures

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Introduction

With the rapid development of computer and information technology, animation is an interesting object to be used for learning and educational purposes. Some research used animation as a tool to improve learning in many domains such as in the school and professional training, in the medical area (in laparoscopic surgery), in the military area, in the aviation area (for example: in air traffic controller, radar operators, and the use of weather displays), in the industrial area (animations as an instructional materials or guidance work when human interacts with a complex system; animation can also reduce error, and enhance safety and comfort in vehicle safety system), in the public safety information (animation or simulation could be used to explain information about how the action should be in an emergency conditions or if natural disasters happen e.g. nuclear reactor explosion or tsunami).

In newspapers or magazines, static pictures showing dynamic processes are used more and more to describe natural disasters or catastrophe of nuclear reactor. For example in Figure 1 and Figure 2 colours are used, with arrow and cueing too. However people generally still found those pictures are too complex or difficult to understand. Usually, the designing process of such pictures is realized intuitively, without any scientific theoretical background.

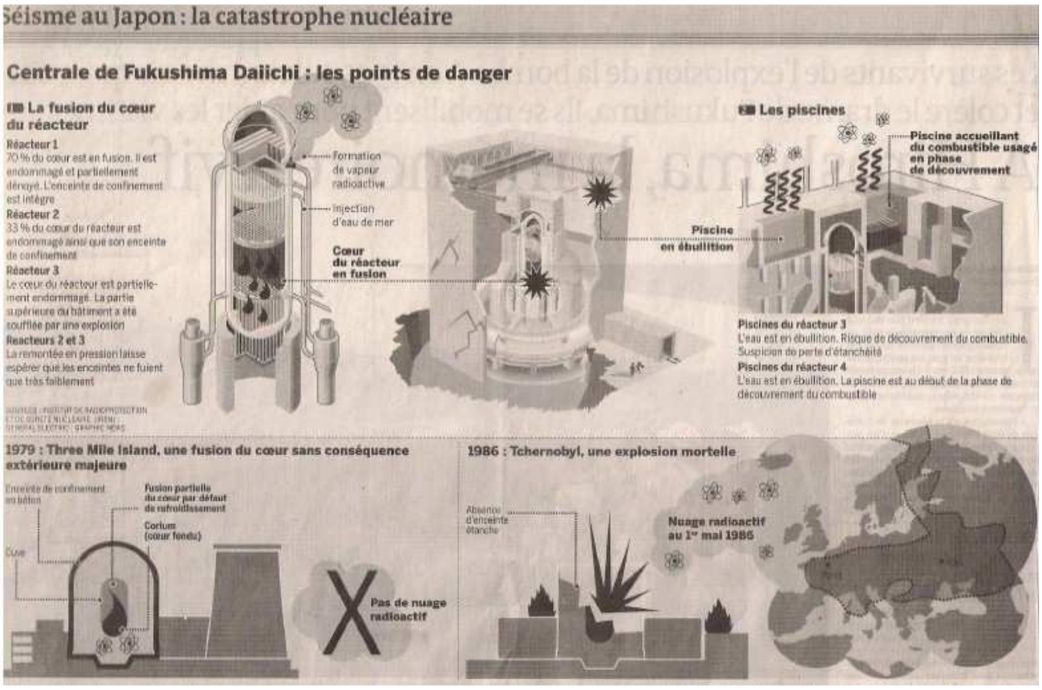


Figure 1. Examples of information in pictures and word without colour cueing (source: Le Monde, 18 March 2011)

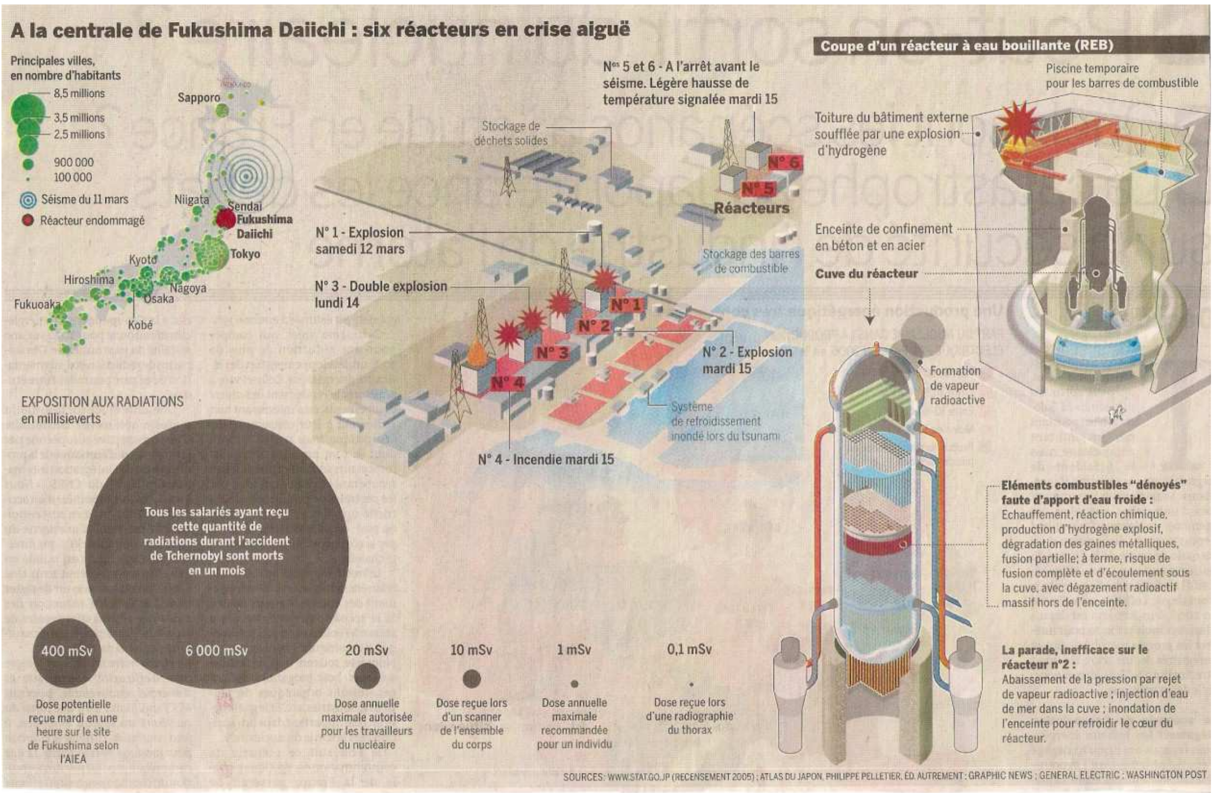


Figure 2. Examples of information in pictures and word with colour cueing (source: Le Monde, 22 March 2011)

Especially in educational purposes, the use of animation instead of the static pictures of school books is increasing continuously. Many researchers and educational practitioners have believed that animation would facilitate learning. Many studies claimed that students learning from animated graphics outperformed those learning from static graphics, for example in Yarden &Yarden, 2010; Huk, 2006, Höffler and Leutner, 2007).

Animations are often considered as superior to static pictures, because they provide more realistic presentation of complex dynamic process, for example: how human lungs works (Mayer,1997, 1999, 2010), how an astronomic phenomenon works (Narayanan and Hegarty, 2002), how machines work (Hegarty and Kriz, 2008), how pumps work (Mayer, 2008), and how piano mechanism works (Boucheix and Lowe, 2008, 2010). However, animation is not always superior to static picture in learning. Some results showed that compared to static pictures animations may even be less effective (Bétrancourt, 2005; Hegarty, Kriz, & Cate, 2003; Tversky-Morrison & Bétrancourt, 2002) and research often failed to find systematic benefits of using animation over static picture (Lowe, 1999, 2004; Morrison and Tversky, 2001).

Several explanations have been proposed for the lack of effectiveness of animations. One of them related to the fact that animations change over time, since animations are transient and they cannot be inspected and re-inspected like the way of static can. The information in animations may be fleeting and hard to process (Tversky, Morrison & Bétrancourt, 2002). Animations are often complex, and it is possible that animations may distract attention and interfere with deeper processing, consequently affecting comprehension negatively. By presenting too much information at the same time, extraneous cognitive load could increase and lead to poorer learning (Sweller, 2005; Sweller & Chandler, 1994).

Although the effect of animation in learning is still controversial, animation became popular, because they are more realistic, interesting, attractive, and motivating (Moreno, 2005). Animations are more realistic for showing change; they can demonstrate in action the systems to be taught.

Until now researchers still explore learning with animation, searching the best way to present animation in order to find “models” or principles of presentation able to help learners for the understanding of complex dynamic system.

In the present work, we examine how to improve the animations that could affect positively learner’s mental models of complex mechanical systems. As a way to overcome difficulties posed by animation processing, we use segmentation and cueing technique in four experiments. The main goal of these techniques is to attract attention on important events or units of the animation and to help parsing activities of the animations. By evaluating the effects of segmentation and cueing in animation processing, we attempt to explain how learners understand a complex system as a result of viewing animations and the criteria necessary for successful learning.

For further, we divided our dissertation into two major parts, a theoretical part followed by an empirical part.

Part one (Theory Part) consists in five chapters. **Chapter 1** is focused on a review about the results of previous research on learning with animation. After giving a definition of animation, we tried to answer the question of whether an animation is better or not than static pictures, and why animation sometimes cannot help comprehension in learning. In **Chapter 2,** we described number of studies which were focused on how to improve learning with animation. We analyzed cueing technique. Besides cueing, we also analyzed segmentation studies which explained why segmentation could be used to improve learning with animation. Interactivity with the display and user-control are also included as an alternative way to improve learning with animation. Another way to improve animation processing could be to use presentation speed variations. Animations can show changes over time at different speed level which enables researchers to manipulate the amount of information that should be processed per unit of time.

Moreover we reported studies which were designed to give to learners explicit strategies for learning from animations. In the end of Chapter 2, a few studies which compared 2D and 3D in animation presentation are analyzed. Using 3D animation refers to the addition of perspective cues to give the impression of depth. Few empirical studies focused on the impact of 3D visualization on learning, however the results were inconsistent.

In **Chapter 3**, we review the main cognitive processing models such as the multimedia model (Mayer, 2001, 2010), the cognitive load model (Sweller, 2005), the mental model theory (Johnson Laird, 1983), and the animation processing model (APM) (Lowe and Boucheix, 2008). **Chapter 4** was focused on spatial ability and prior knowledge which could influence learning with animation. In **chapter 5**, the question was: what do we know about on line processing of animation? Stories of an eye tracking technique were reviewed, and we introduced the new software that we used in the fourth experiment. This software called *ocnoscere*, could be used with dynamic area of interest.

Part two (Empirical Part) is composed of five chapters (6, 7, 8, 9 and 10) with four experiments. In **Chapter 6,** we defined the research question and our goal of research. In c**hapter 7**, the first experiment is presented: ***Chaining two types of visual formats in order to foster inferences in building high quality mental models.*** In **Chapter 8** the second experiment is exposed: ***The use of visual 'tokens' in cueing temporal features of series of static pictures in dynamic process presentation.*** In order to further investigate the results of experiment 1 and 2, we designed experiment 3, presented in **Chapter 9:** ***Cueing temporal features of series of static segmented presentation of dynamic processes followed by an un-cued animation***. Experiment 4 in **Chapter 10,** is focused on a ***comparison of 2D/3D presentation with an eye tracking analysis***. The effect of 2D and 3D in learning complex animation as another alternative of cueing was examined.

Through these experiments, we adapted and elaborated methods to analyse learning from animation processes. In **chapter 11**, we collected all our experiment results and discussed them. Finally, we presented the conclusion of all experiments and gave recommendation for effective learning from a complex animation based on our experiments results.

PART

1

THEORITICAL

CHAPTER 1 *LEARNING WITH ANIMATION*

In this chapter, we will review some basic questions about animation. What is an animation, what are the difficulties of animation processing, compared with a series of static pictures? These studies pose serious challenges to the human information processing system of dynamic picture.

1.1. What is an Animation?

Animation is one of the most attractive forms of pictorial presentations. In the broad definition of animation used by Betrancourt & Tversky (2000) animation is considered to be "series of frames so each frame appears as an alternation of the previous one". According to Mayer and Moreno (2002) animation refers to a simulated motion picture depicting movement of drawn (or simulated) objects.

The view point of three possible effects of animated pictures on the learners was developed by Schnotz and Rasch (2008). The first effect is, the "facilitating effect". Animation can facilitate the construction of a dynamic mental model, mainly by preventing the learners from having to be engaged in a demanding mental simulation. The second effect is, the "enabling effect". Animation could make possible the comprehension of a dynamic system. The third effect is a negative effect, and is called the "inhibiting effect". Animated pictures can inhibit the learners from mentally animating the dynamic phenomenon.

With the rapid development of computer and information technology, animation is an interesting object to be used in learning complex dynamic system. For understanding dynamic systems, three components of dynamics which change over time were identified by Lowe (2004): *the transformation component* involves form changes in size, shape, color, and general appearance of objects, *the translation component* refers to moving objects from one place to another, and *the transition component* involves the appearance or disappearance of objects.

1.2. Is an animation better than static pictures?

From a technological perspective, the differences between static and animated pictures with regard to the physical nature of the representation may well be clear. However, from a psychological perspective, they are fundamentally linked in terms of the human information processing capacities that operate upon them both during learning (Schnotz and Lowe, 2008).

Static pictures should be used to display static content, because they lead to the construction of a static mental model; whereas animated pictures should be used to display dynamic content, because they lead to the construction of dynamic mental model. On the one hand, static pictures are not limited to supporting the construction of static mental models; they can also be the basis for constructing dynamic mental models (Hegarty, 1992). For instance, in representing chemical changing at molecular level, students who were expected to learn any kind of change in matter or motion using static visuals had to visualize those changes using static information, whereas when learning from dynamic visuals, the corresponding changes were apparent (Ardac and Akaygun, 2005). On the other hand, animations are not limited to supporting the construction of dynamic mental models, but they can also lead to the construction of static mental models. Animations can provide learners with explicit dynamic information that is unavailable in the static pictures (Lowe, 2003).

Hegarty, Kriz, & Cate, (2003), using the toilet cistern system, showed that multiple static illustrations of the main phases of the functioning of the system, which showed how the device works, improved mental representation and increased comprehension compared to animated presentations. They examined the effects of computer animations and mental animation on people's comprehension of a relatively complex mechanical system, the flushing cistern (Figure 1.1). The most common use of a flushing cistern is to flush toilets. There were two main subsystems that illustrate different functions, *a flushing mechanism* that flushes water into the toilet bowl and *a refilling mechanism* that refills the tank for the next use.

Their participants studied an animation of the flushing cistern in motion and a series of three static diagrams showing the configuration of machine components at the different key stages of the flushing cisterns in operation (Figure 1.2).

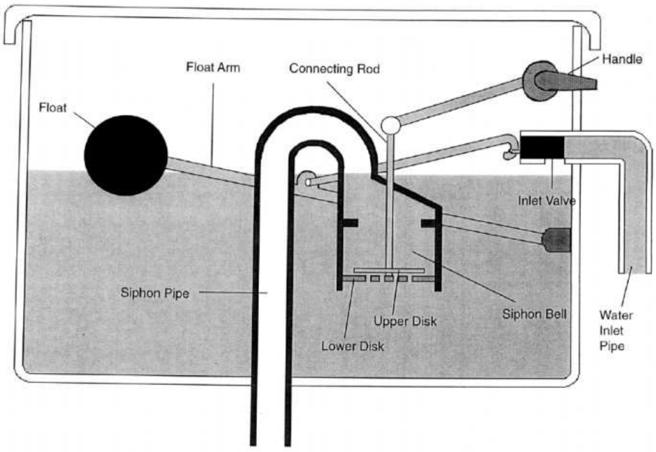


Figure 1.1. Static diagram of the flushing cistern systems (Hegarty, Kriz & Cate 2003)

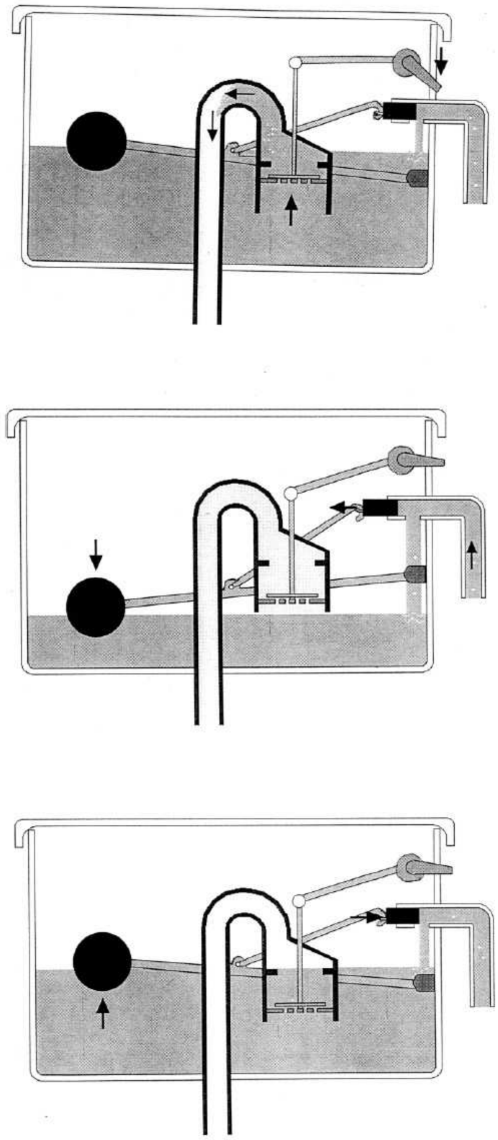


Figure 1.2. Diagram of three phases in the operation of the flushing cistern system (Hegarty, Kriz & Cate 2003)

There was no evidence showing that animation led to superior understanding of dynamic processes compared to the series of static diagrams, although subjects' understanding of the system was improved by viewing both static and animation (Hegarty, Kriz & Cate 2003). Result showed that a series of three static diagrams which showed how the device works, improved mental representation and increased comprehension compared to an animation of the flushing cistern in motion. When presented with static diagrams people have to mentally animate how the system works, whereas when viewing an animation they merely have to perceive how the system works. Mental animation depends also on prior knowledge of mechanics as well as spatial ability (Hegarty, 2004).

In another study, Schnotz & Rasch (2005) compared learning from an animated display (manipulation and simulation) and learning from static pictures of the phenomena related to the earth's rotation (Figure 1.3 and Figure 1.4). In these figures, the earth was depicted as a sphere viewed from the North Pole that rotates in a space where different locations are associated with different states of time.

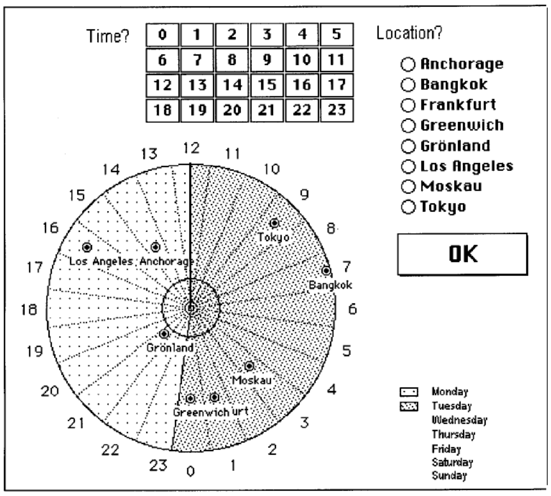


Figure 1.3. Example of a manipulation picture (Schnotz & Rasch, 2005)

Figure 1.3 shows the earth with time zones seen from the North Pole. It can be manipulated by the learner who can define specific day-times for specific cities. After pressing an "OK-button", the earth moves to the corresponding time state. Because the manipulation of the picture enables learners to investigate a high number of different time states, which would not be possible on the static picture, this type of picture is called a manipulation picture and is assumed to have an enabling function.

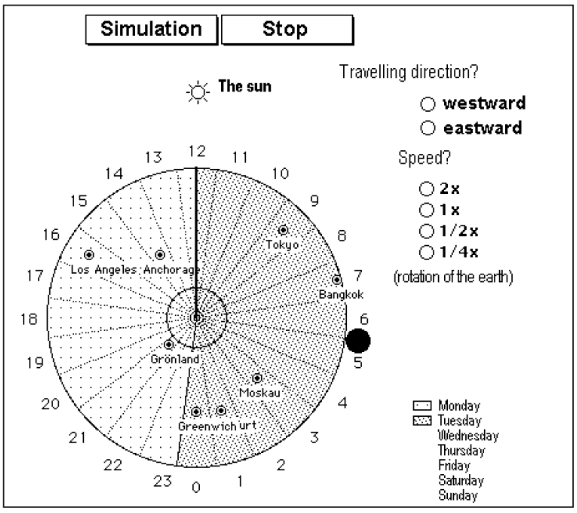


Figure 1.4. Example of a simulation picture (Schnotz & Rasch, 2005)

Figure 1.4 shows a simulation the earth's rotation. Learners can select different ways that a traveller can circumnavigate around the earth (symbolized by a black dot moving in a western or eastern direction, with different travelling speed depending on the learner's choice). After pressing the simulation button, the earth starts rotating and the traveller's dot starts moving on the rotating earth. This type of picture is called a simulation picture, and it is assumed to have a facilitation function (Schnotz & Rasch, 2005). It might be much easier for students to observe the rotation of the earth and the movement of an object in a simulation picture than to perform the corresponding mental simulations on their own with only a static picture (Lowe, 1999; Sims & Hegarty, 1997). Their study was designed to analyze how the assumed functions of animations affect cognitive processing and learning results.

Compared to static pictures, animated pictures can enlarge the set of possible cognitive processes and allow learners to perform more processing than they would be able to perform with static pictures. This was the enabling function of animations. They also stated that animation can trigger dynamic cognitive schemas that make specific cognitive processes easier. They called it the facilitating function of the animations. Result showed that the manipulation pictures had an enabling function for learners with high cognitive ability and high prior knowledge (high learning prerequisites) and simulation pictures had a facilitating function for learners with low ability and low prior knowledge (low learning prerequisites). Learners who had high learning prerequisites, spent more time on animated pictures than on static pictures. Learners who had low learning prerequisites, spent less time on animated pictures than on static pictures. Therefore, the results supported the assumption that the enabling function of animations applies to learners with higher learning prerequisites and the facilitating function applies to learners with lower learning prerequisites.

Boucheix and Schneider (2009) investigated how learners comprehend static and animated presentations in learning dynamic mechanical systems. They used material about a three pulley system from a presentation on a computer screen. The experiment investigated the effect of static *vs.* animated presentations on comprehension. There were four experimental conditions: (1) a single static frame, (2) five integrated sequential static frames, (3) five sequential independent static frames, and (4) an animation. The explanatory text was composed of eight paragraphs in which the configuration of the three pulley system and its functioning was explained (Figure 1.5). The integrated-sequential-static presentation was composed of five frames which depicted the five main steps of the functioning of the three-pulley system presented on the same screen (Figure 1.6).

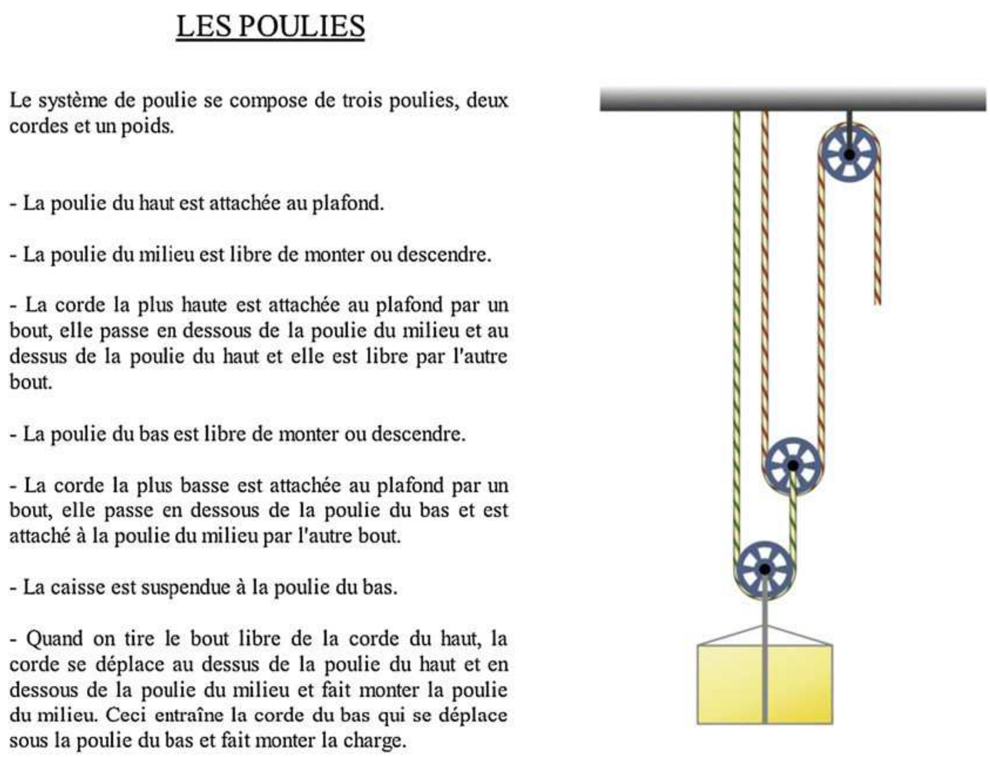


Figure 1.5. Explanatory text and static-frame presentation (Boucheix and Schneider, 2009)

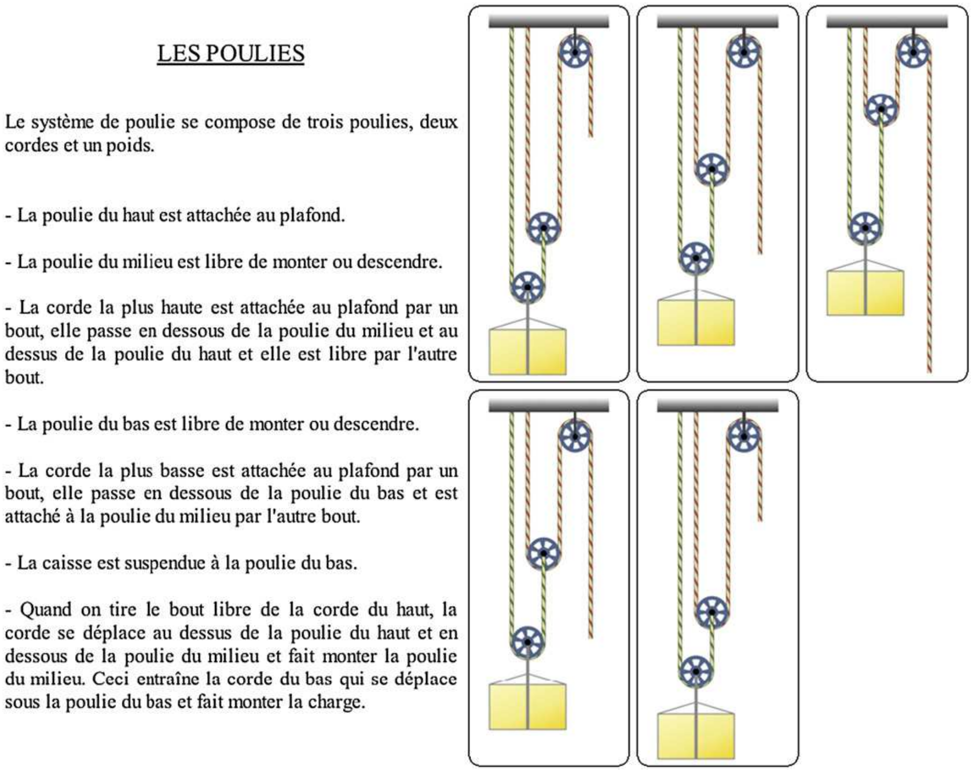


Figure 1.6. Explanatory text and integrated-sequential-static-frame presentation (Boucheix and Schneider, 2009)

Comprehension was measured with a written test that consists of three comprehension indicators, namely configuration, local kinematics and functional mental model (Narayanan and Hegarty, 1998).

The experiment showed that an animation had better results than a single static picture or sequential independent series of static frames. However the animation was not better than the presentation of the simultaneous static frames (figure 1.6.) for the building of a functional mental model of the pulley mechanical system.

Using animations or static pictures could result in differential effects depending on the type of cognitive process involved for learning process. For animation, one might debate that they help in mentally visualizing a process or procedure, resulting in a reduction of cognitive load compared to a situation in which the process has to be reconstructed from a series of static pictures. In static pictures often more or less abstract signalling cues like arrows or highlighting could also be integrated with the pictorial information (Höffler and Leutner, 2007).

Bétrancourt, Dillenbourg & Clavien (2008) reported an experimental study investigating the effects of two features: the continuity of the information flow (animation *versus* series of static graphics) and the permanence of critical snapshots from the animation. The material was about a dynamic phenomenon which described and depicted the formation of lightning storm. This material was used in previous experiments (Mayer, Heiser, & Lonn, 2001; Mayer & Chandler, 2001). In that experiment, they investigated the effect of sequentially displaying snapshots of the main steps of a process while presentation runs. They called "permanence" as the previous steps are permanently displayed on the screen, instead as being moved by the succeeding steps. The original material which included sixteen pictures was reduced to eight according to the model of lightning formation: (1) air rises, (2) water condenses, (3) water and crystals fall, (4) wind is dragged downward, (5) negative charges fall to the bottom of the cloud, (6) the leaders meet, (7) negative charges rush down, and (8) positive charges rush up. The animation stopped after each segment, and the learner controlled the execution as explained by Mayer and Chandler (2001).

Figure 1.7 describes four versions of the material. In the animated presentation without permanence, the animation moves to the centre of the screen. For the animated presentation with permanence, the animation moves to the centre of the screen and when a new sequence begins, the last frame of the previous sequence is placed at the top of the screen.

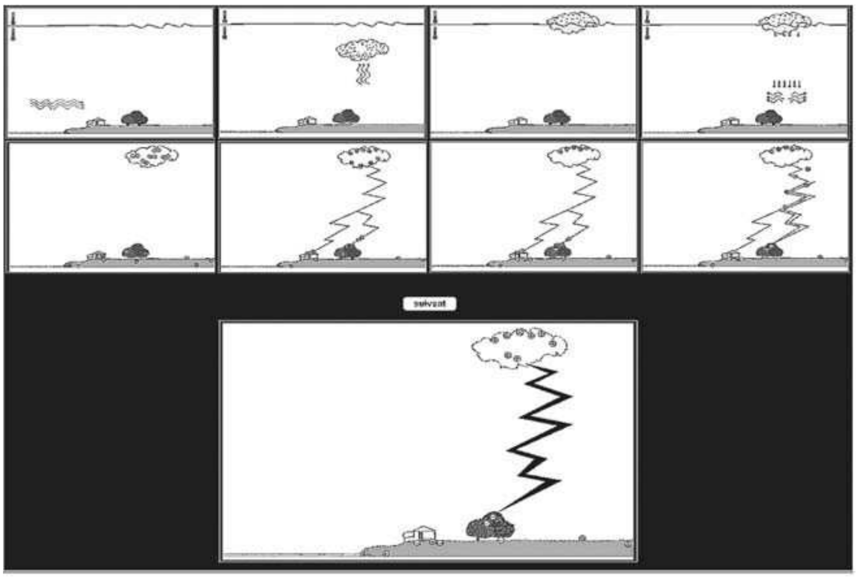


Figure 1.7. Screen display in the last step of the presentation in the condition with permanence. In the conditions without permanence, the upper part of the screen remained empty (Bétrancourt, Dillenbourg & Clavien, 2008)

Result showed that animation was more effective for comprehension than static graphics, but this superiority was mainly due to the study time taken by learners in the animated condition. The learners in the animation conditions had better performance to transfer questions than the learners in the static graphics conditions. These results suggest that learners in the animation conditions elaborated a 'runnable' mental model of the dynamic systems in Mayer's opinion (1989). Moreover, for the retention test which measured surface level meaning of the material, the learners got higher retention score, which means that even the surface structure was better memorized with the animation than with the static snapshots. The learners in the static conditions showed significantly higher difficulties in studying the material than the learners in the animation conditions. Result showed that displaying static snapshots depicting the previous step along with the animation did not act as an effective cognitive aid.

From studies presented above, Höffler and Leutner (2007) stated that animations are significantly better than static presentations when they are representational, that is, the information to be learned related to the motion, trajectory or change over time depicted by the animation.

According to Bétrancourt & Tversky (2000), a well-designed animated display can be expected to facilitate drawing inferences more than remembering information. Animation can lead to cognitive overload. Moreover, if users are not convinced of the necessity of the animation, they might simply not use it.

1.3. Why animation sometimes cannot help understanding of dynamic content?

There are several advantages of using animation in learning systems; (a) animation can help the viewer to build to relevant internal representation, because animation delivers information with an analogical point of view, (Schnotz & Bannet, 2003), (b) animations can be easily adapted to depict dynamic information involving changes over time, because of the similarity in relation with time (Tversky, Morrison, & Betrancourt, 2002), and (c) animations can be adapted for presenting continuous phenomenon, because the learner is not required to infer how phenomena change from one step to the next step (Betrancourt & Tversky, 2000).

However, as the previous review of literature showed, there are also disadvantages of animation. The information in animations may be too rapid and too transient and hard to process because animations change over time, they cannot be inspected and re-inspected as the way static diagrams can (Tversky, Morrison, & Betrancourt, 2002). Animations are often complex, and learners may not know which are the important features to pay attention to and to process.

Moreover, animations can result in specific processing difficulties. The passing of animation imposes information extraction difficulties (Schnotz, 2001). The main source of processing difficulties is the speed at which events are viewed and the number of animated units that could appear within a limited time (Lightner, 2001). Schnotz (2001) described two unexpected effect of animation. The first is the "overwhelming" effect, which happens when too much information is provided to the learners to be processed (Lowe, 2004), the second effect is the "underwhelming" effect. The presence of animation may excuse learners from mentally animating the iconic information, learners fail to process deeply the iconic information from an animation ("learners as viewers", Schnotz, 2001). The "underwhelming" effect happens when learners are guided by the dynamic presentation and do not really get involved in its comprehension because of the complexity of the elements and interactions involved (Lowe, 2003). For learners who are novices in the depicted domain, animations also seem to impose specific difficulties, because with relevant information, they deliver also non relevant visuo-spatial information.

1.4. Should animation be used to improve learning?

Animation can be used to support the visualization of a dynamic phenomenon, when it is not easily observable in real space and time scales (e.g., circulatory system and weather maps), and when the real phenomenon is practically impossible to realize in learning situations. When phenomena are not spontaneously understood in the scientific domain (e.g., the fact that objects of the same volume and different weights fall at the same speed, or the trajectory of falling objects from moving objects), animation can help to visualize those problems.
Animation can also be used to explore and comprehend the phenomenon. In this case the animation becomes a simulation that is used in a discovery-learning approach (Betrancourt, 2005).

Similarly, Hennest et al, (2006) reported that when a process is needed, which involves 'motion', or in complex laboratory experiments that require expensive materials, dangerous or high risk, virtual experiments, it appears that simulation animation is very helpful in reducing errors and in staying focused on particular abstract concepts.

1.5. How does animation affect learning?

Despite of the mixed results found with an animation compared with static pictures (Tversky and Morrison, 2002; Hegarty, 2004; Höffler and Leutner, 2007), many studies have shown examples of unique contributions of animation in promoting conceptual learning, for example in chemistry and biology. For instance in learning cellular biology and molecular processes, the students who viewed a three-dimensional animation of molecular processes were doing significantly better in the follow up test than the group that had not viewed the animation (individual study of text material without using animations), (McClean, Johnson, Rogers, Daniels, Reber, & Slator, 2005). In another study, students who learned about PCR (Polymerase Chain Reaction) by using animation (Figure 1.8) had an advantage in understanding the biotechnological method than students who learned using images (Yarden and Yarden, 2010).

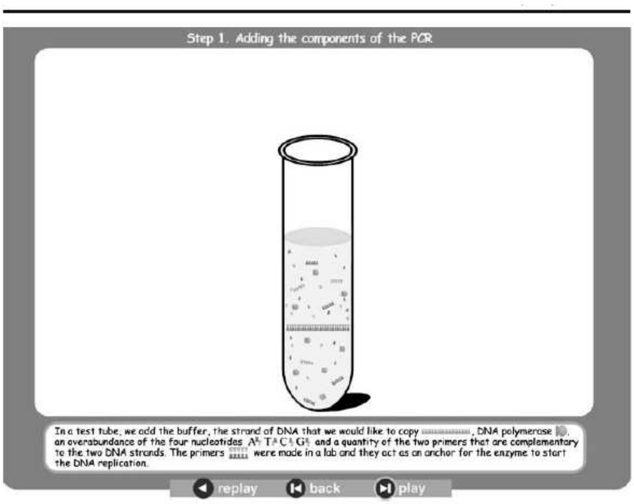


Figure 1.8. A sample frame from the PCR animation (Yarden and Yarden, 2010)

In their experiment they compared the PCR animation with the PCR Cards which contains five cards, each card showing a picture taken directly from the PCR animation and introducing a central stage of the PCR method. The pictures on the cards were accompanied by the same written text that appeared on the relevant frame in the animation.

The PCR animation showed the procedure of increasing a desired DNA (*Deoxyribo Nucleic Acid*) fragment was a test tube using PCR. The subsequent phases of the animation consisted in a demonstration of temperature changes: heating, cooling, and then raising the temperature. It included written texts which appear at the bottom of each screen. The written texts included

images and symbols that explained what was being shown in the animation and what concept each symbol represented.
Result showed that students who learned about the PCR using the cards faced difficulties in understanding the mechanistic aspects in the PCR method. In the contrast, the use of the animation gave the students an advantage in understanding of those aspects.

Ardac and Akaygun (2005) also examined the effectiveness of static versus dynamic visuals to represent chemical change at molecular level. The results also indicated significantly higher performance for students who used dynamic visuals compared with those who used static visuals. Reiber (1990) stated that, dynamic displays can be particularly effective when learning contents that require the students to visualize motion. According to Ardac and Akaygun (2005), one major advantage of dynamic visuals in chemical (when presenting molecular representation) was that they enable a display of collective motion of particles, which can be effective in conveying a complete mental image of changes in matter. On the other hand, when using static visuals, students have to build a complete image by mentally elaborating on what they see on the screen.

From the results of the meta-analysis by Hoffler and Leutner (2007), there was evidence that animations are specifically superior to static pictures when the depicted motion in the animation explicitly refers to the topic to be learned, that is to say when the visualization plays a representational role.
Berney & Bétrancourt (2010) in a fine grain meta-analysis also supported the beneficial effects of the use of animated display for learning dynamic phenomena. Considering the cognitive dimension, three out of the four processes (remember, understand and apply) were beneficial. Students who studied a dynamic phenomenon from astronomy, informatics, or mathematics material performed better. The learner's spatial abilities level positively correlated with the overall efficiency effect, which suggested that the more animation is found efficient, the more the learner's spatial abilities account for it.

CHAPTER 2 *HOW TO IMPROVE LEARNING WITH ANIMATION*

Based on the review in chapter 1, we know that learning from animation could results in processing difficulties. These difficulties could come from perceptual constraints, attentional and working memory resources limitations. Learners must pay attention to simultaneous changes in the display. Learners have also to build a “runable mental model” while they are learning from the animation. They have to maintain each element of the system in working memory (Bétrancourt & Tversky, 2000). In order to improve learning with animation, there are several issues to help learners in selecting, organizing, and integrating information from learning with animation. In this section, we will review recent researches about the main features used to improve of animation presentation: cueing, segmentation and sequentiality, interactivity and users’ control, speed variation, strategy giving, 2D/3D presentation. We referred to previous experimental results.

2.1. Cueing animation

Cueing is the manipulation of visuo-spatial characteristics, but not the content, of instructional materials in order to help learners in selecting relevant information, organizing, and integrating the information into a coherent representation (De Koning, Tabbers, Rikers, and Paas, 2009). Cueing can be intended to direct learners’ attention to essential elements of the visual representation, for instance by increasing the luminance of specific objects in a visual display (De Koning *et al* 2007) or by changing a word’s font style to boldface in a text (Mautone and Mayer, 2001), or by enlarging or highlighting information in text and pictures (Tvesky *et al*. 2008).

Based on Mayer’s theory of multimedia learning, De Koning et al (2009) proposed a framework for classifying different functions of cueing. Three functions of cueing might be related to distinct perceptual and cognitive effects: (1) guiding learners’ attention to specific locations for extracting essential information, (2) emphasizing the major topics and organizing information, and (3) integrating individual elements into a coherent representation. Since several years, numbers of studies were conducted about the effect of cueing in learning from animation. Numerous studies about cueing paradigms in visual attention research have been conducted. The effect of cueing in animations and their characteristics have been summarized by De Koning et al., (2009). Table 2.1 summarized the results of the effect of cueing on learning from animation.

Table 2.1. Summary of the studies on cueing in animations and their characteristics

Author	Modality of cueing	Type of cueing	Function of cueing	Positive cueing-effect on learning
Large et al., (1998)	Visual	Text	Guiding attention Highlighting relations (between representation)	No
Mautone and Mayer (2001)	Visual and verbal	Arrows Intonation Color	Guiding attention Emphasizing organization Highlighting relations (within representation)	Only verbal cueing
Huk et al.,(2003)	Visual	Color Arrows Text (written)	Guiding attention Highlighting relations (between representation)	Yes
Boucheix and Guignard (2005)	Visual and verbal	Color Arrows Text (written)	Guiding attention Highlighting relations (within representation)	Yes (combination verbal/visual cueing)
Lowe and Boucheix (2007)	Visual	Movement Color	Guiding attention Highlighting relations (within representation)	Yes
Kriz and Hegarty (2007)	Visual	Arrows	Guiding attention Highlighting relations (within representation)	No, but cues guide attention
De Koning et al., (2007)	Visual	Luminance contrast (spotlight effect)	Guiding attention	Yes
Moreno (2007)	Visual	Color	Emphasizing relations (between representation)	No

Van Oostendorp and Beijersbergen (2007)	Visual	Color	Guiding attention Highlighting relations (between representation)	No
Fischer and Schwan (2008)	Visual	Movement (speeding up) Arrow	Guiding attention	Yes, especially movement
Fischer et al., (2008)	Visual	Movement (speeding up)	Guiding attention	Yes
De Koning et al., (2009)	Visual	Luminance contrast (spotlight effect)	Guiding attention	No, but cues guide attention
Boucheix and Schneider (2009)	Visual	Arrow	Guiding attention Highlighting relations (within representation)	No

Source: De Koning,B.B., Tabbers, H.K., Rikers, R.M.J.P., & Paas,F. (2009)

As stated by De Koning, et al., (2009) in table 2.1, cueing could be effective in directing learners' attention to cued information, but it is not always effective in improving learning outcomes.

Huk, Steinke, & Floto (2003) studied the added value of non-verbal cues. They used colour, arrows, and text (written) to serve two functions of cueing. Arrows and colour are used as guiding attention to essential information in the animation and text (written) as highlighting relation between representations (animation and narration). Result showed that cueing resulted in a better learning performance.

The function of the visual colour as a guiding attention to essential information also was used in learning from complex animation, investigated by De Koning et al., (2007). The study stated that the effectiveness of visual cues is dependent on the complexity of the instructional animation and only enhances learning if learners need cues to guide them in building a coherent representation (De Koning.,Tabbers.,Rikers., & Paas, 2009).

Fischer and Schwan (2008) examined whether manipulating the speed of animation and arrow cues were equally effective to guiding attention in an animation. Result showed that manipulating the speed of animation was a significantly more effective cue and more salient for guiding attention and improving learning than having learners study the animation with locally focused arrow cues. Lowe and Boucheix (2007) also supported the idea that the dynamic cueing effectively stimulated learners to cognitively process the content more deeply.

However, cueing does not enhance learning performance in a systematic way (Large et al, 1998; Kriz and Hegarty, 2007; Moreno, 2007; Van Oostendorp and Beijersbergen, 2007; and De Koning, 2009). Results failed to prove a positive effect of cueing on learning performance.

Large et al., (1998) studied the effects of cueing in an animation of the cardiovascular system by using text caption. Result showed that adding captions to the animation did not enhance learners' understanding of the system. This lack of effect may be due to the labels that were included in the animation. These labels made the captions redundant, which consequently added very little to understanding of the content (De Koning, Tabbers, Rikers, and Paas, 2009).

Van Oostendorp and Beijersbergen (2007) studied the effects of cueing by highlighting a part of the animation and simultaneously placing a dot before the sentence referring to that part of the animation. Result showed that the cueing condition did not succeed in relating specific concepts in the text with the corresponding elements in the visualization. It means that cueing might not have been specific enough to facilitate processing in that study.

Similarly, Moreno (2007) reported that cueing did not improve learning performance. Cueing may have forced the learners to spatially split their visual attention between the animation and the highlighted labels that were presented side-by-side, and it may have interfered with the learning process.

These mixed findings of the cueing studies paralleled those found by Kriz and Hegarty (2007) and De Koning (2009). Both of those series of studies demonstrated that visual cues effectively guide attention, but do not necessarily enhanced comprehension of the content (e.g no effect on learning outcome).

From the table of the review of the 13 studies, we tried to make a summary based on the type of cueing (visuo spatial, cueing, and temporal cueing) and the effect of each type of cueing (attention guiding and learning outcome) in table 2.2.

Table 2.2. Summary of the 13 studies reviewed by De Koning et al., (2009) about cueing

Type of cueing		Effect of cueing				Total
		Attention guiding		Learning outcomes		
		Positive	Negative	Positive	Negative	
Visuo spatial (2)	Text	1	-		1	1
	Verbal	1	-	1	-	1
Cueing (9)	Arrow	4	-	2	2	4
	Color	3	-	1	2	3
	Luminance	2	-	1	1	2
Temporal cueing (2)	Movement	2	-	2	-	2
	Total	13	0	7	6	13

Table 2.2 shows that on the 13 studies, it possible to conclude that cueing is useful for directing learner's attention to specific parts of an animation, but cueing does not always enhance comprehension. Only temporal cueing seems to be efficient for improving comprehension.

Moreover, Ozcelik, Karakus, Kursun, and Cagiltay (2009) investigated how color coding affects multimedia learning. Their experiment examined the effects of color coding on learning outcomes and examined why learners performed better when using color-coded material as compared to monochrome material. The material experiment focused on chemical synapses (Figure 2.1 and Figure 2.2).

Information about chemical synapses was presented on a first computer screen and on a second screen was presented two different formats: the conventional format, included an illustration and a text which were physically separated from each other, and the color-coded format, which was identical to the conventional format except that all of the field specific terminology was presented in the same colour within the text and in the illustration.

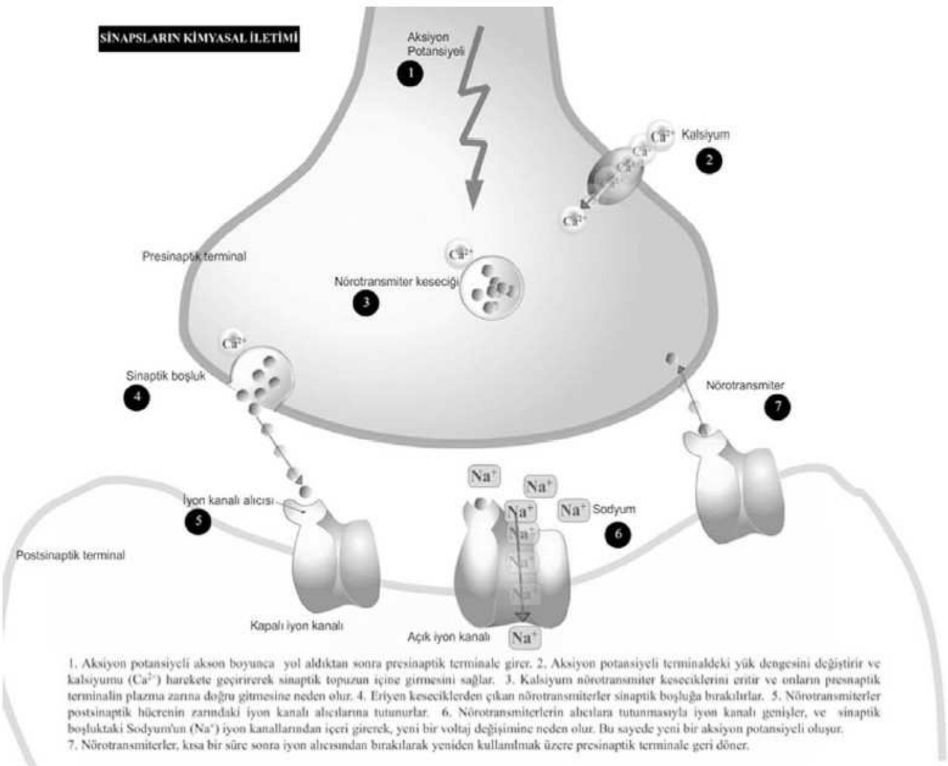


Figure 2.1. The Conventional format
Ozcelik, Karakus, Kursun, and Cagiltay (2009)

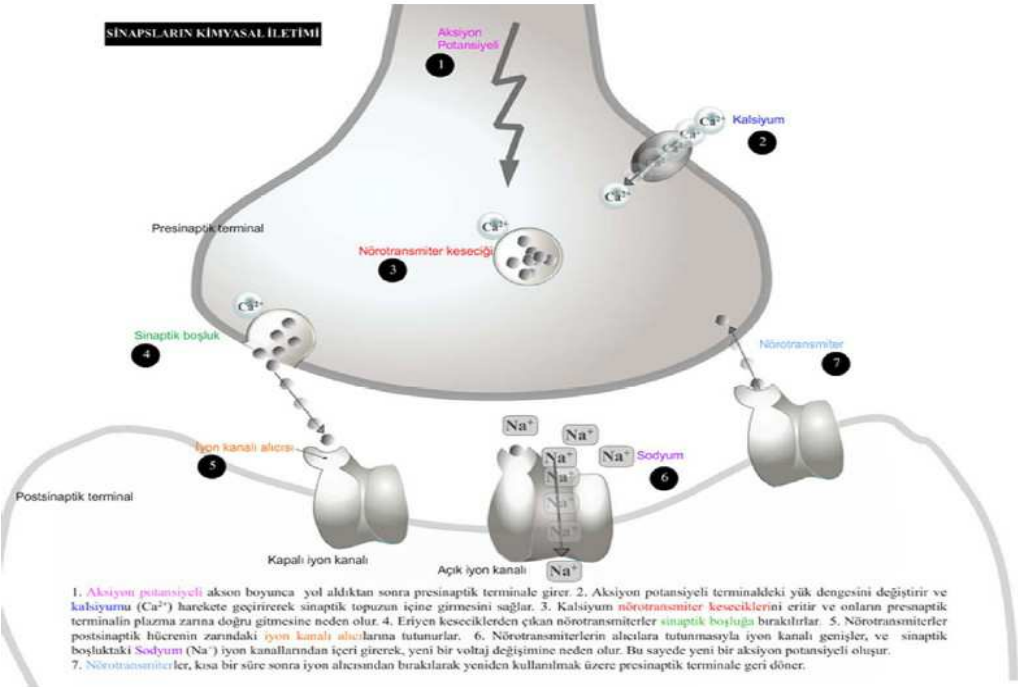


Figure 2.2. Color coded format
Ozcelik, Karakus, Kursun, and Cagiltay (2009)

Result showed that it was easier for learners to find relevant information and to integrate verbal and pictorial information in colour-coded format than in the monochrome format.

Recently, evidence of positive effects of cueing was also found for animation by Boucheix and Lowe (2010), they emphasized that cueing supported the construction of a mental model of causal chains. The colorization occurred in temporal correspondence with the events in the animation and different colours were used for different subsystems. The visual cues conveyed information spatially and temporally (Figure 2.3).

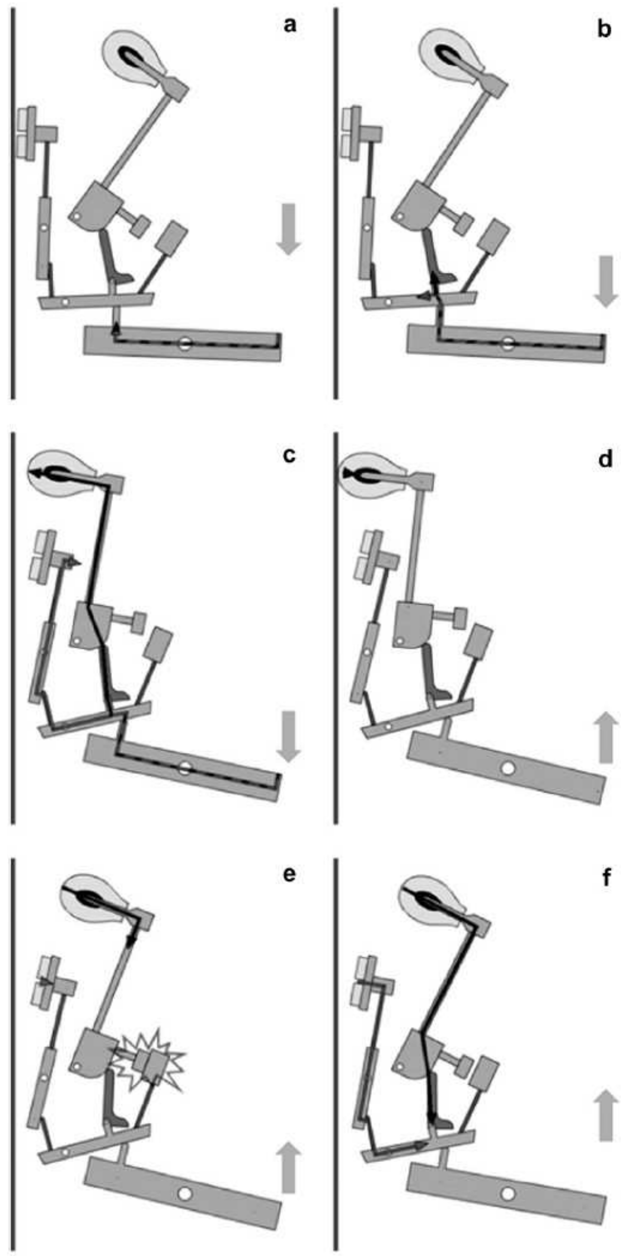


Figure.2.3. The spreading-colour cue (a-f) (Boucheix & Lowe, 2010)

During an animated presentation, learners need to coordinate spatial and temporal aspects of their visual exploration of the relevant contents. Result showed that spreading-colour cues resulted in better targeting of attention to thematically relevant aspects and in higher comprehension scores than arrow cues or no cues.

De Koning, Tabbers, Rikers, and Paas (2010) also reported about cued system in learning complex animation of the human cardiovascular system. De Koning et al., (2010) used shading of all elements in the animation except the cued sub system (Figure 2.4).

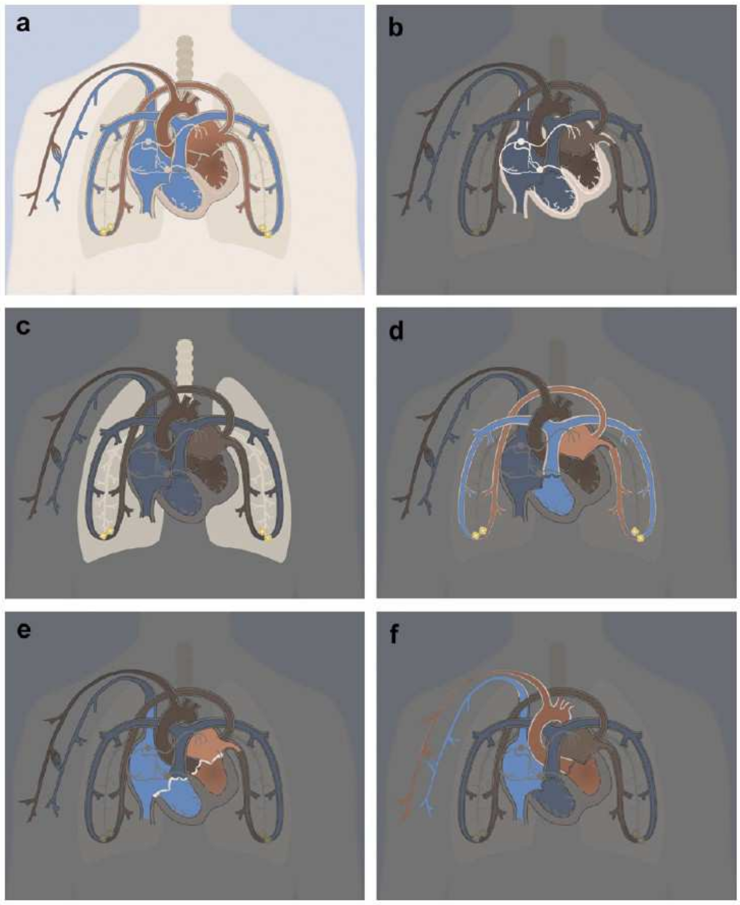


Figure 2.4. (a) the non-cued animation and (b-f) the cued subsystem. Note: a = no-cueing, b = electrical system, c= circulatory system, d = pulmonary circulation, e = valves system, f = system circulation (De Koning, Tabbers, Rikers, and Paas, 2010)

The presentation depicted dynamics of the system's five main subsystems: circulatory system, electrical system, pulmonary circulation, systemic circulation, and valves system. The animation was presented without accompanying written or verbal descriptions, no pauses, and learners had no opportunity to control the animation. The cued sub system covers only the valves system.

This visual contrast allows the cued subsystem to become more salient because its perceptual attributes differ from those of the rest of the animation. As in the Boucheix & Lowe (2010), result showed that learners looked more often and longer at cued than at non-cued parts.

2.2. Segmented and Sequential presentation of dynamic processing

In everyday activity, people segment on going activity into meaningful temporal parts. In particularly, people focus on the events that make up everyday life on the timescale.

Zacks and Tversky (2001) stated that an event is "a segment of time at a given location that is perceived to have a beginning and an end". Sometimes it is difficult to say where or when one event ends and another begins.
To identify the boundaries between event parts, Zacks (2004) suggested two possibilities: (1) people could break an activity into parts based on typical sensory characteristics. For an example, the lights in a cinema dim before the movie. A dimming cinema light might cause a person to notice that an event has ended solely, because it is a salient sensory change, even if that person had never been to the cinema and did not know that the film was begun. This kind of identifying event boundaries is *bottom-up*, because the sensory characteristics serve as direct cues to the structure of event, (2) people might identify event boundaries by recognizing the activity in progress and parsing it in terms of one's knowledge of that activity. For an example, a rocket leaving the ground might be perceived as an event boundary because the viewer recognizes a rocket launch as beginning with a countdown, followed by take-off. Segmenting based on knowledge structures is *top-down,* because sensory characteristics are interpreted in terms of representations of the activity that are abstracted from physical stimulus and depend on prior knowledge.

The animation supply information in perceptual form. However, the information in animations may be fleeting and hard to process (Tversky et al., 2002). One of the factors that cause learners difficulties in focusing their attention (and do not get involved on essential information in animation) is the complexity of the elements and interactions involved which can lead to an overwhelming feed of information and that is not adapted for individual learning speed. The possibility to overcome that problem is offered by segmenting a diagram into different frames and by presenting them progressively.
The segmenting principle explains why learners' understanding of materials is better when they can control the advance of the presentation from one segment to the next rather than viewing a continuous presentation (Mayer and Chandler 2001; Mayer et al., 2003; Moreno 2007).
The segmenting effect is entrusted to practical use in condition where the intrinsic load of the materials is so high that learners do not have enough cognitive resources for the essential processing of the content. When the materials are segmented and the learners are allowed control over when to advance to the next segment, learners may be able to effectively dispose

this load over the amount of time they need to process the materials (Lee et al., 2006; Mayer 2005; Mayer and Moreno, in press). Table 2.3 provides an overview of studies on segmentation and sequential presentations with their characteristics.

Table 2.3. Summary of studies on Segmented and Sequential presentations

Author	Instructional domain	Type of segmentation	Function of segmentation	Effect of segmentation
Mayer and Chandler (2001)	The formation of lightening	Segmentation *vs.* animation	Guiding attention Emphasizing relation	+ effect for attention guiding + effect on learning outcome
Rebetez, Bétrancourt, and Sangin (2009)	Astronomical animation	Static picture sequence vs. animation sequence with and without snapshots.	Guiding attention Construct mental model	+ effect for attention guiding - effect on learning outcome
Betrancourt and Tversky (2000)	Map of library	Sequential vs. static display	Guided explanation Construct mental model	+ effect for attention guiding + effect on learning outcome
Betrancourt, Dillenbourg, and Montarnal (2003)	Financial sheet analysis	Sequential non interactive Sequential interactive Static condition	Guiding attention Construction mental model	+ effect for attention guiding + effect on learning outcome
Arguel and Jamet (2009)	Technique of first aid	Video *vs.* static picture *vs.* combination video and static picture	Guiding attention	+ effect of attention guiding + effect on learning outcome especially for mixed format.

A first research about segmentation in multimedia presentation has been conducted by Mayer and Chandler (2001). The material was composed of slides which conveyed information about the formation of lightning in sixteen steps (and characterized by strong causal effects between the different steps). After a segment was presented, the learner could get the next segment by clicking on the continue button. In this way, the learner could fully process one segment before moving on to the next. The segmenting principle is that people learn better when a narrated animation is presented in learner-paced segments rather than as a continuous unit.

Mayer and Chandler (2001) presented learners with an animated multimedia explanation on the formation of lightning and compared it to a version that was previewed in smaller segments under the control of the learners. They hypothesized that previewing the animation segment by segment would eliminate the cognitive load compared to a presentation of the animation as a whole, because learners would have more time to mentally organize the information of each segment separately.
It would allow the learners to fully understand each components of the causal chain before moving into the next one, thereby eliminating the risk of cognitive overload and supporting deep understanding, as measured by a transfer test.

In addition to segmentation, there is another method called as sequential display which presents a diagram in a progressively growing way. This method offers a guided explanation of a complex system as it conveys “the organization and inherent logic of the instruction, just as a teacher draws a schema on a blackboard in a carefully chosen sequence” (Bétrancourt, Dillenbourg, Montarnal, 2003).

Bétrancourt and Tversky (2000) reported an experiment about a map of library with the sequential display. The result showed that the sequential display was not significantly better than the static display. It should be especially effective for complex graphics whose structure is not immediately apparent. The sequential presentation of a complex graphic display can guide the construction of a mental model of the graphic as well as facilitate memory for it.
As stated by model of Mayer (2001), presenting information sequentially make it possible for the user to construct first a local mental model, which can be later integrated in a coherent mental model.

Another experiment study on sequential displays was designed by Bétrancourt, Dillenbourg and Montarnal (2003). The learning material was a multimedia lesson on financial analysis, containing graphics and corresponding text. The lesson explained about how to transform an accounting balance sheet into a financial sheet. Three experimental conditions were designed: (1) a sequential non-interactive condition, the order in which elements were displayed was defined by the instructional designer and was computer controlled; (2) a sequential interactive condition, the order in which elements were displayed was under learner’s control; (3) a static condition where the elements appeared simultaneously on the screen with all text available in an adjacent window. The results showed that the learners with a sequential interactive condition were fastest, followed by the learners with a static and sequential non-interactive condition.

According to Hidrio and Jamet, (2006), there are several reasons for the benefit of sequential displays as it reduces the cognitive load presenting information gradually, because of the following reasons: there are less visual elements visible at the same time that should make it easier to process the information; the learner needs less time for visual search. Further, a sequential display offers a coherent order of information processing and it can be used to regroup the elements of the same category in time, even if they are spatially distant; a sequential display promotes the construction of a functional mental model when a dynamic system is, presented.

Jamet, Gavota and Quaireau (2008) tested the effect of a sequential display experiment by examining the effects of adding color to the elements which are presented (Figure 2.5). The experimental conditions were designed across four different types of presentation: a static presentation of the diagram, a sequential presentation of the same diagram, a static presentation in which an element of the diagram was highlighted and a sequential presentation with the relevant element highlighted.

Their hypothesis stated that the best performances of all conditions should be in the sequential display with highlighted elements of the diagram. Sequential presentation should facilitate selection by limiting the number of information presented on the screen. The use of color should capture attention automatically and should facilitate the selection of the relevant element for a contact in processing memory. They concluded that sequentially and salience could be combined to obtain the highest possible level of benefit; it could function as attention guidance on the remembering of multimedia documents.

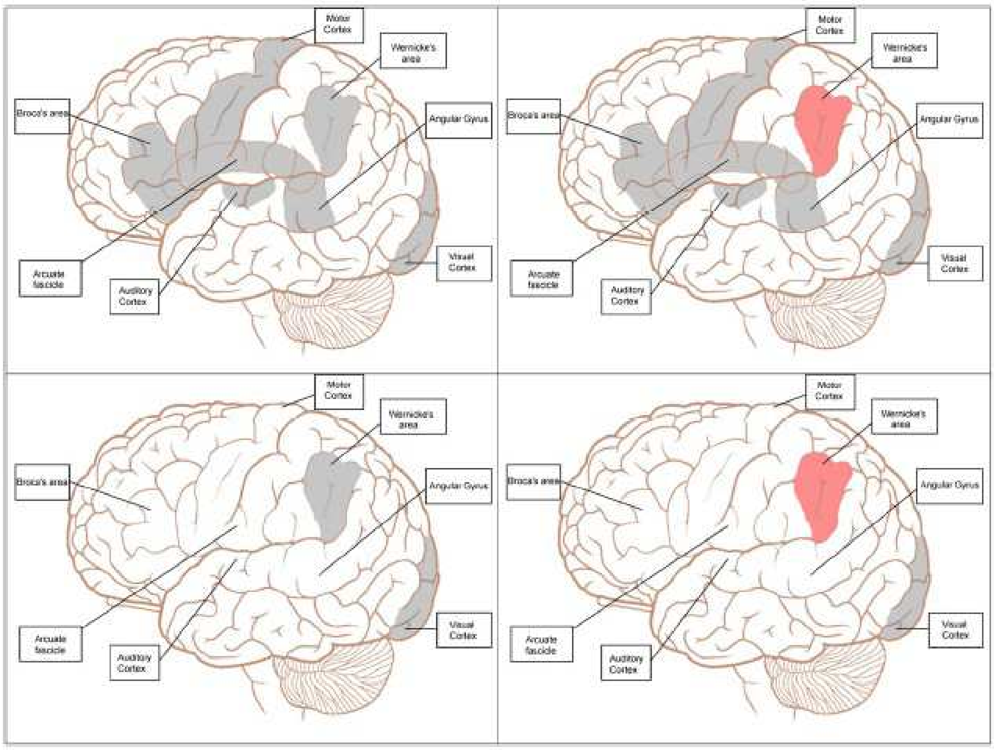


Figure 2.5. Static/non-salient group (top left), static/salient group (top right), sequential/non-salient group (bottom left), and sequential/salient group (bottom right) (Jamet, Gavota Quaireau , 2008)

There are many various recommendations for helping designers to use multimedia presentations with efficiency in various learning environments. To learn content phenomena that change over time, the use of dynamic visualizations such as animations and videos are often designed and used as a way to help understanding and facilitate learning.

The effect of presenting together both a video recording and a series static picture was investigated by Arguel and Jamet (2009). Their experiment was composed of three conditions: (1) video shown alone, (2) static pictures displayed alone, and (3) video plus static pictures. Result showed that participants in the combination format (video plus static pictures) outperformed the two other formats. The alternative mixed format would make better learning outcomes than the video, because the static pictures added to video would limit the transience of some information. The alternative mixed format would be more effective than the static pictures alone, because the static pictures would not contribute enough information to the learners given the relatively complex nature of the materials to be learnt. The addition of a series static picture into videos could be beneficial, because the steps depicted in pictures were actually the most important steps of the procedure described (Arguel, and Jamet, 2009). This situation could create an external representation closer to the mental representation expected of learners than offered when only videos are presented. This result was consistent with the congruence principle; a combined material should be more efficient in encouraging learning than visualizations composed only of videos (Tversky et al., 2002).

Moreover, they also investigated the impact of the precise number of supporting pictures required (low vs. high frequency) and their appearance type (static vs. dynamic) to find out how best to present the static pictures. The opposite result was found: the dynamic presentation of pictures was superior to the static pictures, and fewer pictures presentation (low frequency) was more beneficial. Based on the cognitive load point of view, showing too many pictures could enhance extraneous load and consequently have a disadvantage effect on learning performance (Arguel & Jamet, 2009).

Based on several studies on segmentation and sequential in Table 2.3, it is possible to conclude that there are not many researches on segmentation and sequentially. Segmented and sequential modalities could result in positive effect in both of attention guiding and learning outcome.

2.3. Interactivity and Control

Interactivity principle could be defined by the idea that people learn better from a multimedia instruction when they are able to control both order and pace of the presentation (Mayer et al. 2003). To facilitate learning, interactivity is a factor that can help to overcome the difficulties of perception and comprehension. Stopping, starting, and replaying an animation can allow re-inspection, focusing on specific parts and actions. Animations in which close-up, zooming, and control of speed to focus on specific parts and actions are more likely to facilitate perception and comprehension (Tversky, et al., 2002).

Types of control in animation can be divided into (a) control of slides and (b) a control of the speed. Research by Mayer and Chandler (2001) tested the effect of learner control in animation which depicted the process of lightning formation accompanied by an explanative narration. A control device allowed breaking down the animation in short sequence by pausing automatically. The learner had control to run each animated sequence to the next sequence. Result showed that giving learners the controlled animation provided better transfer performance than giving learners the continuous animation.

On the other hand, some studies have found contradictory result: Lowe (1999, 2004) found that studying an interactive animation on weather maps did not result in the use the interactive features in a very effective way. Schnotz et al (1999) also found that interactive animation were less effective than static pictures and showed that the interactive features only enhanced cognitive load and did not support learning.

Boucheix and Guignard (2005) investigated, with children from 10 to 12 years old, the effect on comprehension of the illustration format, animated and static and the function of signalling and rhythm of the information giving. They used multimedia document explaining the function of a simple mechanical system “how a gearing system works”: rotation of the gear wheels, direction, relative speed of each wheel and mechanical effect. The functioning of gearing systems is described in Figure 2.6.

They focused on two factors on comprehension of animated and static conditions, signalling cues (presence *vs* absence), and rhythms of information giving (fixed fast, fixed slow, and self-controlled). The results showed a positive effect of animations when compared to static version and an effect of the signalling cues for the explanations concerning the system.
The positive effect of animations especially appeared in the self-controlled. The presence of signalling cues was only efficient in guiding the attention focus on the learner in the self-controlled modality.

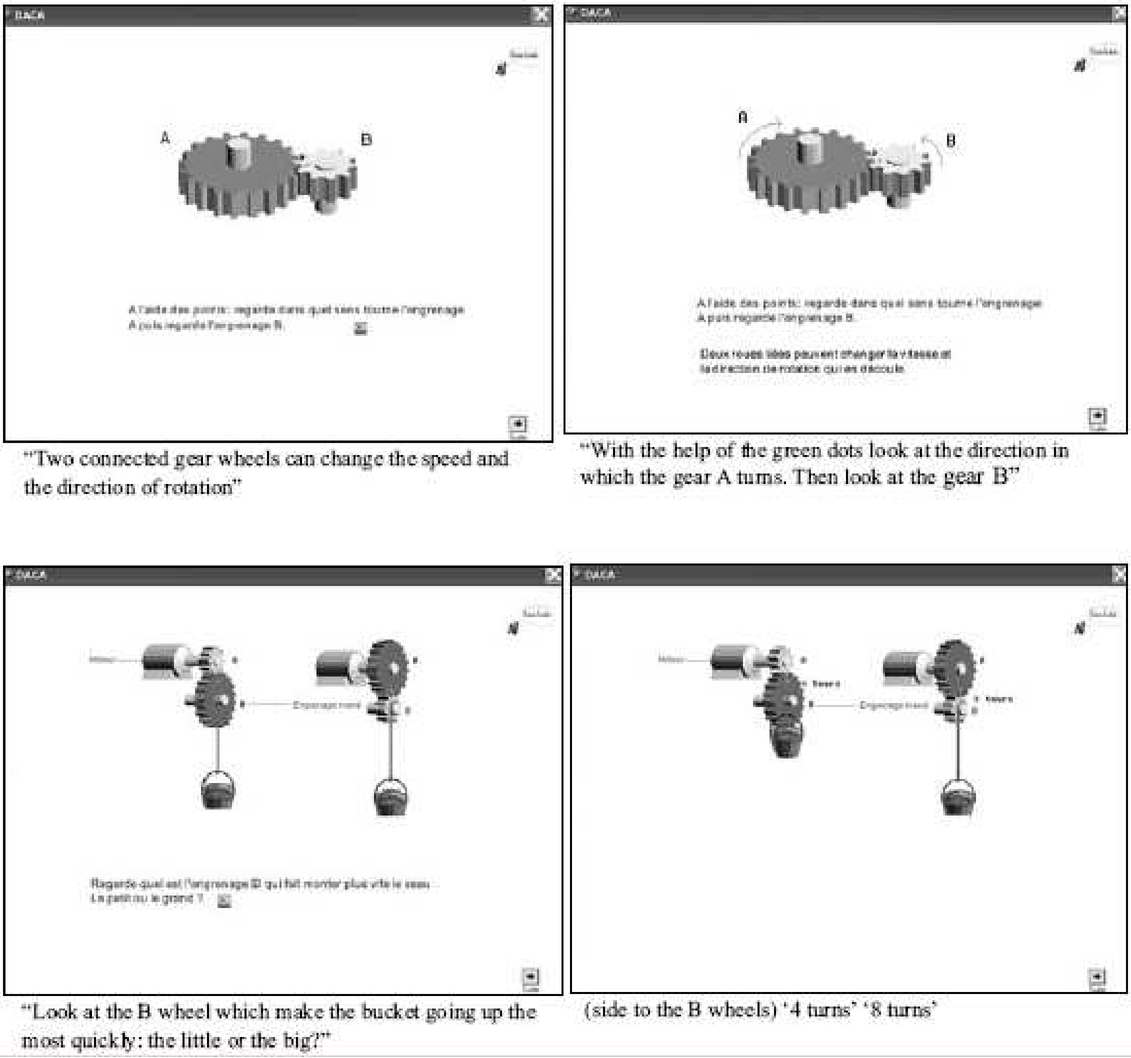


Figure 2.6. The functioning of gearing systems (Boucheix and Guignard, 2005)

2.4. Speed variation

Cueing may be especially effective for improving animation-based learning if animation has high complexity. Recent studies investigating the effects of an animation's presentation speed on learning have indicated that changed speeds may improve understanding (Fischer, Lowe & Schwan, 2008; Meyer, Rasch, & Schnotz, 2010).

Animation can show change over time at different speed levels, which enables researchers to manipulate the amount of information that should be processed per unit of time and thereby the load that is imposed on learners' perceptual and cognitive resources, without removing or adding elements from the content, by increasing or decreasing an animation's speed.

Decreasing presentation speed might make it easier for learners to extract relevant parts and reduces the possibility of only partially processing or missing relevant parts, because learners have more working memory resources available for exploring the animation due to more time and less visual search requirements. A low speed may allow learners to construct a mental representation of local parts which then can be integrated into a coherent mental model (Meyer, Rasch, & Schnotz., 2010). In contrast, increasing presentation speed may force learners to quickly and repeatedly decide which information requires intentional processing as the same amount of information is presented in less time. Consequently, learners may miss or partially process information due to high visual search rate and limited time to relate and integrate current with previous information in order to comprehend the animation, thereby increasing working memory demands (Ayres & Paas, 2007).

Recent study about the role of presentation speed was conducted by De Koning, Tabbers, Rikers, & Paas (2011). In their study they investigated a cued or an un-cued animation about the cardiovascular system (Figure 2.7) at high or a low speed.

It was hypothesized that the high speed animation would require cueing to facilitate the identification of task-relevant elements based on Kriz & Hegarty, (2007). Their result was unexpected: students showed equal performances on comprehension and transfer test irrespective of cueing and the animation's speed. However, the low speed groups invested more mental effort to obtain learning performances than the high speed groups.

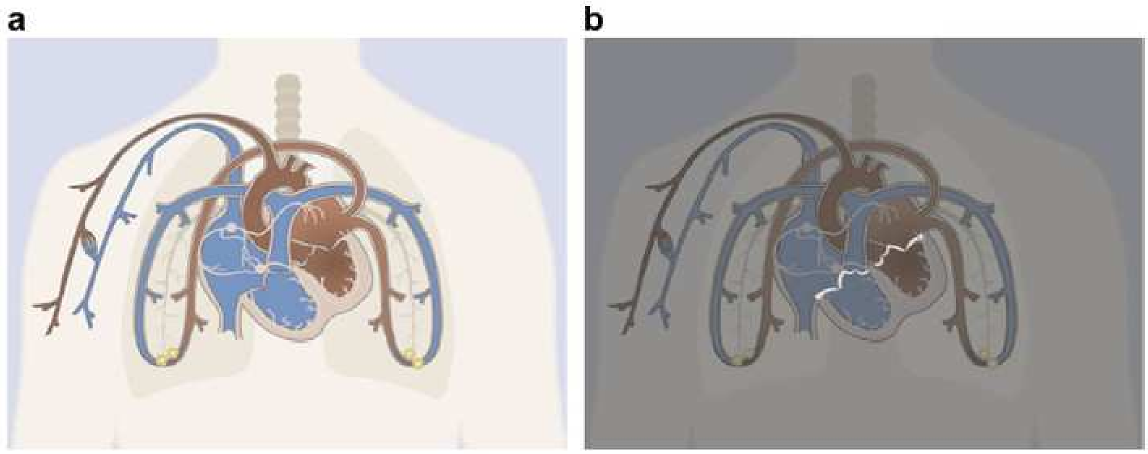


Figure 2.7. Un-cued (a) and the cued (b) animation of the cardiovascular system (De Koning, Tabbers, Rikers, & Paas, 2011)

It might be thought that the low speed group may have developed demands on WM, thereby creating extraneous load (Ayres & Paas, 2007). The additional WM demands may have resulted in higher learning outcomes for the low speed groups than the high speed groups. In order to support participants in learning from complex animation more effectively, more precise knowledge is needed about what perceptual and cognitive effects instructional manipulations such as cueing and presentation speed have on animation processing (De Koning et al., 2011).

Meyer, Rasch, and Schnotz (2010) investigated also the effects of high and low presentation speed of animation. Participants learned about the functioning of a four-stroke engine from an animation with user-controlled presentation speed. Structure of a four-stroke internal-combustion engine which consists of cylinders, valves, pistons, and so on was described on figure 2.8.

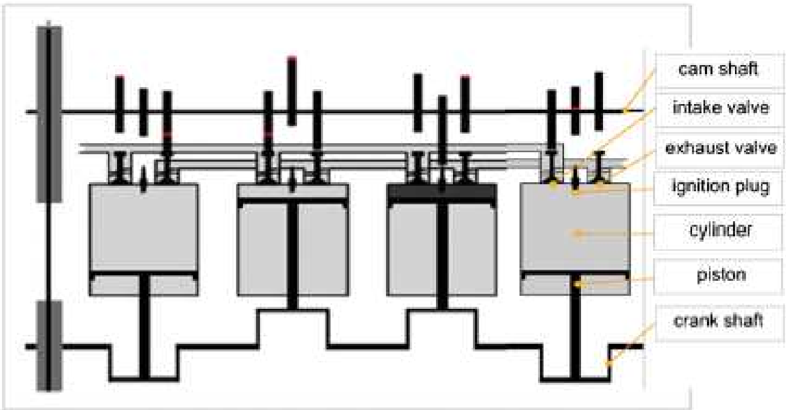


Figure 2.8. Structure of a four-stroke internal-combustion engine (Meyer, Rasch, & Schnotz, 2010)

Results showed that the speed of the animation's presentation had an impact on comprehension of micro and macro level events. The learners preferred a low speed of the animation's presentation to observe micro-level events of animation. The higher the presentation speed, the more macro-events comprehension was achieved compared to the learner's total comprehension score that included both macro and micro event scores. The presentation speed of animation affects the salience of information at different levels and the learning outcomes.

Fischer and Schwan (2010) reported an experiment about the presentation speed of the clock mechanism (normal vs. Fast) with signalling cues on parts of the mechanism.

They hypothesized that both the effects of speed and spatial cueing should have a beneficial effect on learners' understanding of the pendulum clock's mechanism. The version of the operating mechanism of the pendulum clock is shown in figure 2.9.

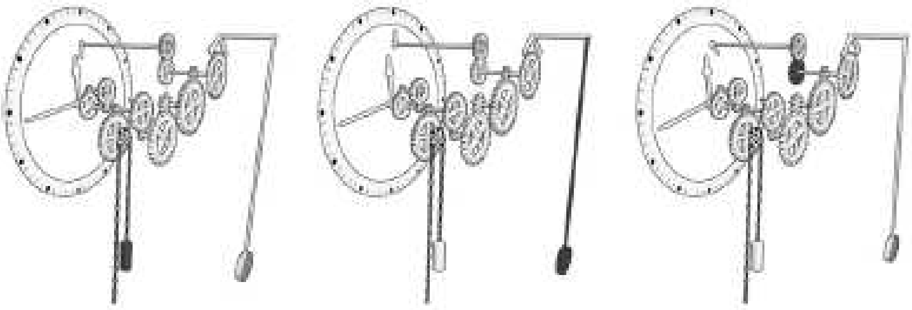


Figure 2.9. Frames showing cued parts (in dark color)-weight, pendulum, and gear (left to right) (Fischer & Schwan, 2010)

The animations displayed the operating mechanism of the clock uncovered and the clock face indicated only by a ring with marks. Cued clock parts were in the weight, the pendulum, and the lower of the two gears.

Contrary to the weight and the pendulum, the gear has no importance for the main functional mechanism of the clock. Cueing was realized by changing colour from original object colour to red colour cyclically. Results showed that understanding of the operating mechanism of the clock was improved in the fast speed condition but there was no interaction of tempo with signalling cues. The use of cueing condition did not interfere with the temporal manipulation. Although explicit cueing was successful for guiding attention to the cued parts of the clock, as proved by the results of the rating scales, however it has no impact on comprehension of the operating mechanism of the pendulum clock.

2.5. Giving a strategy for animation processing

Previous methods used to improve animation processing were mainly bottom-up techniques; however top down techniques could also be interesting. We saw that cues were effective in directing attention to relevant location, but not effective in learning outcomes. This means that cue effect is mainly implicit and maybe more explicit instruction are needed.

Research on learning with animation has not yielded any proven strategies that would enable learners to systematically and successfully approach animations. As translated by Kombartzky, Ploetzner, Schlag, & Metz (2010) based on Streblow and Schiefele (2006), a learning strategy is defined as “(a) a sequence of efficient learning techniques, which (b) are used in a goal-oriented and flexible way, (c) are increasingly automatically processed, but (d) remain consciously applied”. Learning techniques should result in specific internal learning activities such as remembering a piece of information and establishing a relation between two pieces (minimum) of information, or external learning activities such as highlighting and annotating segment of a text (Streblow & Schiefele, 2006). A learning strategy is formed when some of several learning techniques are used in a coordinated and goal-oriented way.

A strategy for learning from animation based on theories of multimedia learning was designed by Kombartzky, Ploetzner, Schlag, & Metz (2010). Two experimental studies were conducted for evaluation of the proposed strategy. The learning material was an animation combined with spoken text which showed with visualizations how honey bees dances in order to communicate to other bees where resources in the environment are located (Figure 2.10).

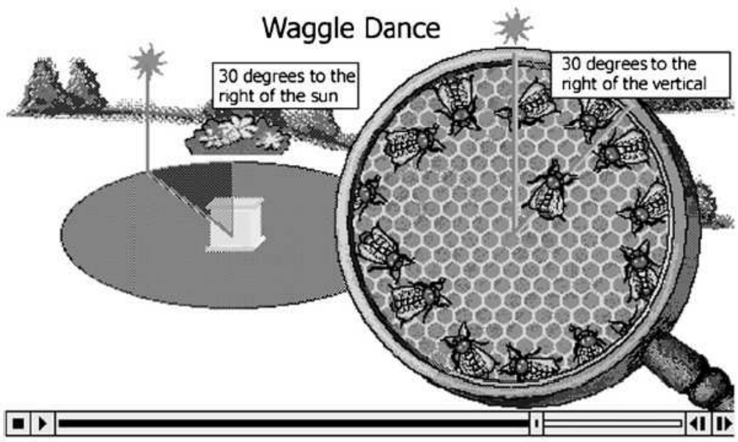


Figure 2.10. A screen shot taken from the animation (Source: Schlag and Ploetzner, 2010)

They proposed couple of hypothesis; (1) it was expected that the students who were encouraged to make use of the learning strategy should learn more successfully than the students who were not encouraged, and (2) it was expected that the students whose use of the strategy was monitored should learn more successfully than those whose use of the strategy was not monitored. Result showed that the students who were encouraged to take advantage of the strategy learned significantly more than the students who were not asked to do. Further, learning was most successful when the students' use of the learning strategy was monitored. In both studies, they demonstrated that the proposed learning strategy is suitable for improving learning from animation. The strategy was more supportive as the levels of knowledge to be acquired became more demanding.

2.6. 2D *versus* 3D format comparisons

Many kinds of disciplines and research groups are concerned about the design and effectiveness of visual display. Statisticians are concerned with how to best design graphs to describe data (Cleveland, 1985). Researchers in human factors and engineering psychology developed and examined principles of design for domains such as process control, aviation, and medicine (Smallman & St. John, 2005; Smallman, St. John, Oonk, & Cowen, 2001; Agus, Bettio, Gobbetti, & Pintore, 2007). In the educational domain, many researchers have developed dynamic interactive visualisations to teach student about scientific processes such as chemical and mechanical interactions or to relate different representations of the same phenomenon (Ainsworth, 2006; Wu, Krajcik, & Soloway, 2001).

Hegarty (2011) explained that visual displays can be categorized into three types based on the relation between the representation and its referent and the complexity of the information represented: (1) *Iconic displays*: visual-spatial displays that represent visual-spatial entities, (2) *Relational displays*: visual-spatial displays that represent abstract relationship, and (3) *Complex displays*: multiple representations, animation, and interactivity. Examples of visual displays are shown in Figure 2.11.

Figure (A) shows a mechanical diagram representing a pulley as circles and figure (B) is a road map. They represents of objects which are visual spatial entities. Iconic displays can show views of their referent that are not visible in real-world viewing (Hegarty, 2011).

Figure (C) is a scatter diagram that shows the relationship between two variables and Figure (D) is Euler's circles represent the premises in a reasoning problem. In these displays, visual and spatial properties represent entities and properties that are not necessarily visible or distributed over space. Figure (E) is a map showing levels of pressure and temperature which is used by meteorologist.

Figure (F) is fMRI data that visualized a map of activity across the brain. Figure (E) and (F) noted a hybrid of iconic and relational displays in displaying non visible properties overlaid on a representation of a visual-spatial entity.

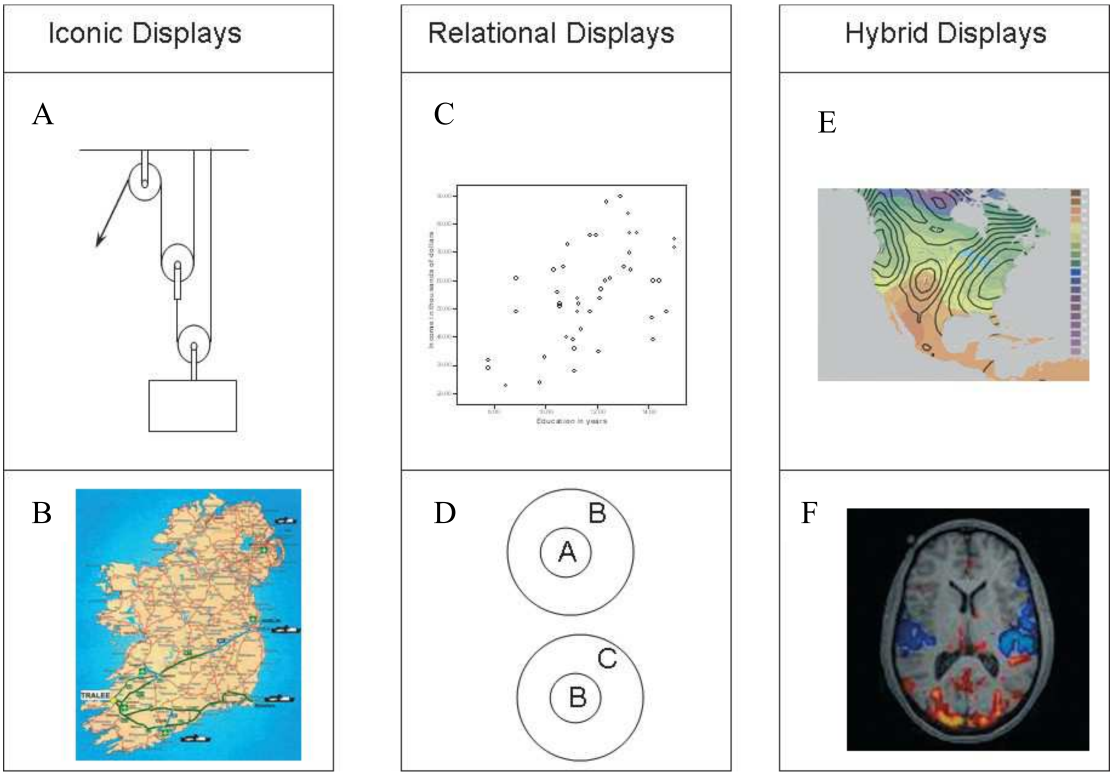


Figure 2.11. Examples of types visual displays: iconic, relational, and hybrid displays (Hegarty, 2011)

There are many other examples of complex displays, such as multiple representations, or animations in which a sequence of images shows the key steps of a dynamic process (Tversky, Morrison, & Betrancourt, 2002), or interactive displays that allow the user to rotate the display in order to view an object from different point of view, or even, three-dimensional data set from different perspectives (Robertson, Czerwinski, Fisher, & Lee, 2009).

Research in the field of comparing two-dimensional (2D) and three-dimensional (3D) visualization on learning are still rare and inconsistent. Researchers are looking for new and better ways to visualize complex system in 2D/3D visualizations. When and how to use 2D or 3D visualizations for specific situation and tasks has been done for example in medical visualization tasks (Agus, Bettio, Gobbetti, & Pintore, 2007), in air traffic controllers (Smallman, St. John, Oonk, & Cowen, 2001), in military objects (Smallman, St. John, Oonk & Cowen, 2000), in geographic terrain (Hollands and Ivanovic, 2002), and in topographic maps (Wiebe, Slykhuis, & Savage, 2005).

In medical visualization system based upon 3D spatial displays provided lots of advantages. The system presented multiple viewers the illusion of seeing virtual objects floating at fixed physical locations. The usage of 3D visualization techniques helped disambiguating complex images, so it was accepted to be a real advantage for immediate understanding and visualization of medical data. Three D displays were able to provide all the depth cues exploited by the human visual system (Agus, Bettio, Gobbetti, & Pintore, 2007).

However, in military research, recent result showed that 3D realistic icons are harder to identify than conventional 2D military symbols (Smallman et al., 2000). Topographic maps research showed that using 3D map did not affect students' ability (positively or negatively) to gather information in earth science education (Wiebe, Slykhuis, and Savage, 2005). However, students who studied stereochemistry stated that they preferred either 2D or 3D representations depending on the focus of the stereochemistry tasks or type of isomers that they were investigating (Pavlinic, Buckley, Davies, and Wright. 2001).

Cockburn and McKenzie (2004) has investigated and evaluated spatial memory in two and three dimensions. They determined whether three dimensional user interfaces better support spatial memory than traditional two dimensional. They reported no significant difference between the effectiveness of spatial memory when using 2D and 3D computer-supported systems, but significantly better performance with 2D than 3D when using the physical systems.

Other research which was focused on the use of 3D models was reported by Huk (2006). He investigated the educational value of 3D visualizations on learning and understanding of cell biology. Result showed that only students with high spatial ability benefited from the presence/absence of 3D models, while students with low spatial ability got fewer points when learning and understanding cell biology. Their data demonstrated that the students with low spatial ability became cognitively overloaded by the presence of 3D models, while high spatial ability students profited from them, as their total cognitive load remained within working memory limits. They concluded that the educational value of interactive 3D models depends on spatial ability. Further experiment that added visual cues in 3D representation formats in an animation in the domain of biological education was conducted by Huk, Steinke, and Floto (2010). They wanted to know whether a 3D representation format combined with visual cues implemented with an animation would foster knowledge acquisition. Their result clearly indicated that the implementation of visual cues enhanced the amount students remembered in experiments. Students' comprehension of the dynamic process in domain of biological education was facilitated by the presence of a 3D representation format under tightly controlled conditions only (restricted system-paced).

From the many results of experiments presented above, it seems still inconsistent and debatable whether the implementation of 3D models can enhance learning processes. We may conclude that both 2D and 3D visualizations are useful for improving learning processes however it depends on each conditions (spatial ability's personal, purposes in learning process). In general, 2D visualizations are good for seeing details of particular part of object, for the tasks that require judging the precise distances and angles objects. Three D visualizations are good for gaining an overview of 3D and also useful to shape understanding because it provides natural depth cues showing the object from varying perspectives.

CHAPTER 3 *COGNITIVE PROCESSING MODELS*

In the domain of multimedia learning and animation comprehension several cognitive model have been recently proposed. The ability to understand or to access to relevant information from the source of learning materials relies on specific cognitive processes. Understanding the role of the learner's cognitive processing is particularly important during learning process. The review presented in this chapter emphasizes the importance of comprehension processes in the design of multimedia information and animation. There are several models which help to understand how the information processing of an individual happens.

3.1 Cognitive theory of multimedia learning

The cognitive theory of multimedia learning developed by Mayer (2001, 2009) is based on three basic assumptions: (1) the dual-channel assumption, (2) the limited capacity assumption, and (3) the active processing assumption.

The dual-channel assumption is the idea that humans have two separated channels for processing visual/pictorial channel and auditory/verbal channel (Baddeley, 1998; Paivio, 1986). Limited capacity assumption states that only limited information can be processed at the same time in each channel (Baddeley, 1998; Sweller, 1999). Active processing assumption is means that learners need to be actively engaged in the learning process, such as selecting relevant material, organizing incoming information, and integrating with existing knowledge. The model of multimedia learning by Mayer (2001) is presented in Figure 3.1

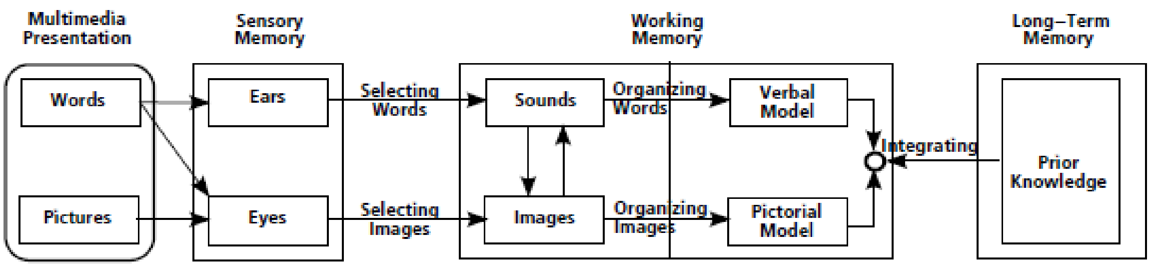


Figure 3.1. Cognitive theory of multimedia learning taken from 'Learning with Animation' pages 34 (Mayer, in Lowe 2008).

Figure 3.1 shows a cognitive model of multimedia learning which consists of three memory stores: sensory memory, working memory, and long-term memory, and five cognitive processes: selecting words, selecting images, organizing words, organizing images, and integrating. Pictures and words stimulate the eyes and the information enters through visual

channel whereas narration stimulates ears and auditory information enters through the auditory channel. The first cognitive activity consists in *selecting* relevant material for further processing.
The *selecting words* arrow indicates that the learner pays attention to some of the sounds in sensory memory. Information is transferred to working memory for the next processing, and if the learner pays attention to some of the images in sensory memory, they are transferred to working memory for the next processing that indicated by the *selecting images* arrow. The second cognitive process is organizing the selected material into coherent representations: the learner mentally organizes the words into a verbal model, point out by *organizing words* arrow and the learner mentally organizes the images into pictorial model, point out by *organizing images* arrow. In the end the two models are integrated with prior knowledge from the long-term memory.

3.2. Cognitive load theory

Every learner has a limited amount of cognitive capacity in any learning condition. There are three processing demands on cognitive capacity: *extraneous processing*, which involves cognitive processing that is not related to the instructional objective but to the external material organization; *essential processing* which involves basic cognitive processing of material that is relevant to the objective of the lesson, and *generative processing*, which involves deep processing of material that is relevant to the objective of the lesson (Mayer, 2001). Sweller, Van Merriënboer & Paas (1998) uses term *intrinsic load* to refer to *essential processing* and *germane load* to refer to *generative processing*. Learning with animation can be particularly challenging when the material is difficult, not familiar, and presented with a fast pace. There are ten ways to overcome challenges to learning with animation proposed by Mayer in Lowe (2008) as listed in table 3.1. Particularly, segmenting principle and signalling principle are most closely related to our experiments presented in the next chapters.

Many researchers have started using the cognitive load theory (CLT) as a relevant model to explore the effects of instructional manipulation of animation. CLT is based on a model of human cognitive architecture which assumes that working memory (WM) is very limited in terms of being able to store and process information (Cowan, 2001; Miller 1956); whereas long-term memory (LTM) has a great capacity, able to store almost unlimited of information.

Table 3.1. Ten ways to overcome challenges to learning with animation (Mayer in Lowe 2008, pages 37)

Reducing Extraneous Overload *Coherence principle*: people learn better when extraneous animation or narration elements are excluded rather than included. *Signaling principle*: people learn better when cues are added that highlight the organization of the animation or narration. *Redundancy principle*: people learn better from animation and narration than from animation, narration, and on-screen text. *Spatial contiguity principle*: people learn better when corresponding elements of the animation and on-screen text are presented near rather than far from each other on the screen. *Temporal contiguity principle:* people learn better when corresponding animation and narration are presented simultaneously rather than successively.
Managing Essensial Overload *Segmenting principle***:** people learn better when a narrated animation is presented in learner-paced segments than as a continuous unit. *Pre-training principle*: people learn better from a narrated animation when they have had training in the names and characteristics of the main concepts. *Modality principle*: people learn better from animation and narration than from animation and on-screen text.
Fostering Generative Processing *Personalization principle*: people learn better when narration is in conversational style rather than formal style. *Voice principle*: people learn better when the narration is spoken in a standard-accented human voice rather than a machine voice or accented human voice.

The comparison of animation and static graphic in previous research seem to have empirical evidence to support the cognitive load theory. Some researchers of the superiority of animation over static pictures may claim that animation decreases cognitive load by providing an external visualization of a process or a procedure that would have to be otherwise mentally reconstructed and visualized by the learner from a series of static pictures. In other words, as Schnotz (2002) it is possible to claim that animation may have an enabling function which could benefit to the learners who cannot mentally reconstruct or visualize a complex dynamic process or procedure from a series of static pictures.

However, research by Hegarty (2004), showed that when an animation is composed of several static pictures and displayed sequentially with each frame being only available for a short period of time; one may argue that processing animated information imposes higher cognitive load due to the temporal limits of working memory.

A study by Amadieu, Mariné, & Laimay (2011) was carried out focusing on the impact of cueing on cognitive load and comprehension of animations which described a dynamic process in a neurobiology domain. Using CLT as theoretical background, their study investigated how cueing may limit the attentional requirements for processing of animations. The cueing method used in their study consists of a zoom on each step of the dynamic system which hides irrelevant information at each step (Figure 3.2 and Figure 3.3).

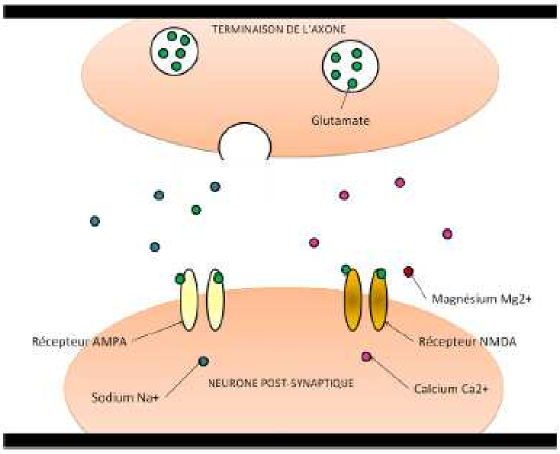


Figure 3.2. Screenshot of a step of the animation without cueing (Amadieu, Mariné, & Laimay, 2011).

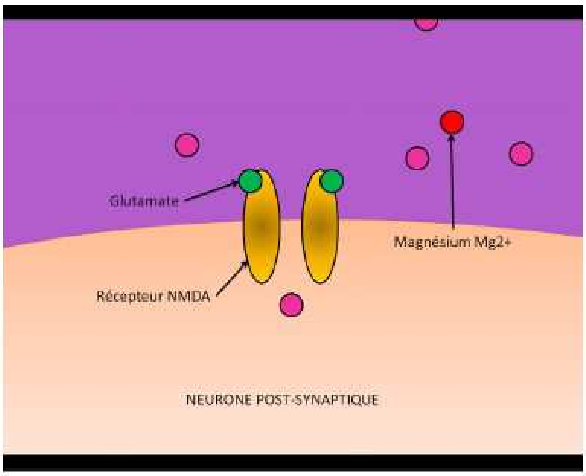


Figure 3.3. Screenshot of a step of the animation with cueing (zoom), (Amadieu, Mariné, & Laimay, 2011).

Figure 3.2 showed the beginning of each step the mechanism of Long Term Potentiation which is a chemical and an electrical phenomenon occurring in synapses mechanisms. A purple square appeared on the relevant region to direct attention to it. Figure 3.3 presents a zoom made on the region displaying only the relevant elements of the steps, hiding the irrelevant elements that were expected to avoid distraction to elements not involved in the step. Result showed that extraneous cognitive load was reduced by cueing after three exposures and a problem solving task showed that the development of a more elaborate mental model was supported by cueing.

3.3. The initial mental model theory

A mental model is an explanation of someone's thought process about how something works in the real world. This is a form of knowledge representation which is constructed from perception, imagination, alternatives and hypotheses, when engaged in a process of mental simulation (Johnson-Laird, 1983).

Mental models are based on a small set of fundamental assumptions, which distinguish them from other proposed representations (mainly propositionals) in the psychology of reasoning (Byrne and Jhonson-Laird, 2009). Each mental model represents a possible representation of the state of the real world in the mind. A mental model represents one possibility, capturing what is common to all the different ways in which the possibility may occur (Johnson-Laird and Byrne, 2002).

Building a mental model from a multimedia presentation requires apprehending a selection of relevant information, organizing the information, and integrating it with prior knowledge (Mayer, 2005; Narayanan & Hegarty, 2002; Schnotz, 2002). The building of a mental model emerges from interaction between information from visual perception and cognitive processes which are guided by general knowledge of each individual (Schnotz, 2002).

According to Johnson-Laird (1993), the building of a mental model starts with the formation of mental images. These images consist in representations characterized by a high level of analogy with the perceptual characteristics of real objects and their spatial relationships. Further an example, the breakdown of mental images in order to achieve mental reasoning leads to recoding of the perceived object into a set of elements, named "tokens", which are meaning full partial structures of the objects. The learner operated changes and mind control on the tokens. The causal relationships way between the various elements of the model are made explicit by the creation of inferences. Inferences processing requires mental "manipulation of simple model composed with "tokens".

The final stage of the model consists in searching for alternative models for purposes of validation of the model built. Johnson-Laird (1998) argued that in order to be controlled in working memory, mental models need to be simplified representations or "small-scale models" of the external world.

Several types of mental models can be built, such as a dynamic mental model. They can be developed as part of learning a complex dynamic system which consists of a continuous time representation of the operation of the system.

3.4. General cognitive processes in visual-spatial displays: the notion of display schema

Hegarty (2011) proposed a general model of diagram processing. The model was adapted from Freedman & Shah (2002) and Pinker (1990).

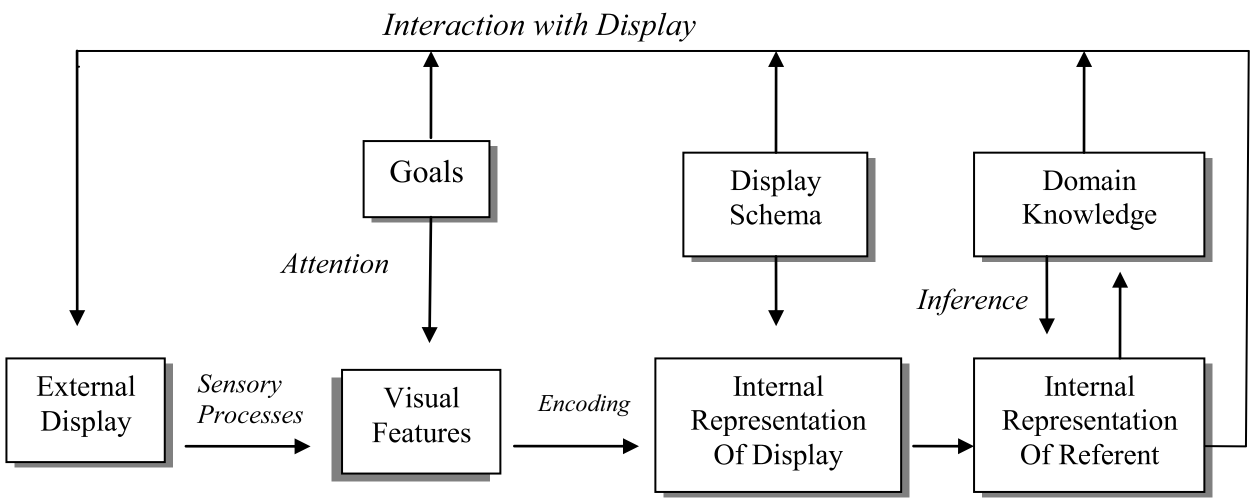


Figure 3.4. Diagram of the different representations (indicated by boxes) and processes (indicated by arrows) involved in understanding a visual display (Hegarty, 2011)

Internal representation of display is constructed by the visual system which apprehend the senses of the features of the display (ex: colour and shape) and by the encoding of these features. The encoding of the features relies on attention, which might be directed by the viewer's goals and expectations or by what is salient in the display. The user of a display has to apply knowledge to build a representation of its referent. The type of knowledge has been discussed mainly in the context of graphs and referred to as a graph schema (Pinker, 1990). However Hegarty (2011) used the more general term of "display schema" (in Figure 3.4). The understanding of a display can fail, if the user's display schema is incomplete.

Comprehension of a display can also include making further inferences based on domain knowledge or visual-spatial processes, so the resulting internal representation comes to contain information that is not presented explicitly in the external display (Hegarty, 2011). However this model is mainly based on previous research using static displays.

3.5. Principles related to the perception of visual displays

There are three principles related to the perception of dynamic displays: (1) The apprehension principle (2) The principle of discriminability and (3) the principle of perceptual organization. The apprehension principle was formulated by Tversky et al. (2002). They explained why animations of physical processes are not more effective than static graphics. The apprehension principle is also related to the use of visual dimensions that are more or less accurately perceived (Cleveland and McGill, 1983) and which avoid or not the use of visual variables that lead to biased judgements. Smallman & St.John (2005) found that a perceptual error could be relevant for a successful rendering of geospatial displays in 3-D. They argued that 3-D displays are not effective for judging absolutes distances because people misperceive distances in depth.

The principle of discriminability comes from basic research in psychophysics. This principle means that visual forms indicating dissimilarity between two variables need to differ by a large amount of information to be perceived as different. The principle of perceptual organization is based on Gestalt principles. This principle determined which elements of displays are grouped and can be compatible or incompatible with the tasks to be carried out with the display (Kosslyn, 2006).

3.5.1. Event Segmentation Theory (EST)

Zacks, Speer, Swallow, Braver, and Reynolds, (2007) argued that perception is a hierarchical process in which sensory information is successively transformed into representations that form the basis for action. Perceptual predictions are representation of states of the world in the near future. In their Event Segmentation Theory (EST), Zacks et al., (2007) explained that Event Segmentation Theory arises from the perceptual processing stream as depicted in Figure 3.5.

The input of a pathway is a set of sensory representations and the output is a set of perceptual predictions. The sensory inputs relate to the information conveyed by the peripheral nervous system to the cortex. Sensory inputs are transformed by perceptual processing to produce multimodal representations with rich semantic content, encoding information such as object identity and location, motion trajectories, and the identities and attitude of other people.

An event model is a representation of “what is happening now” which is robust to transient variability in the sensory input.

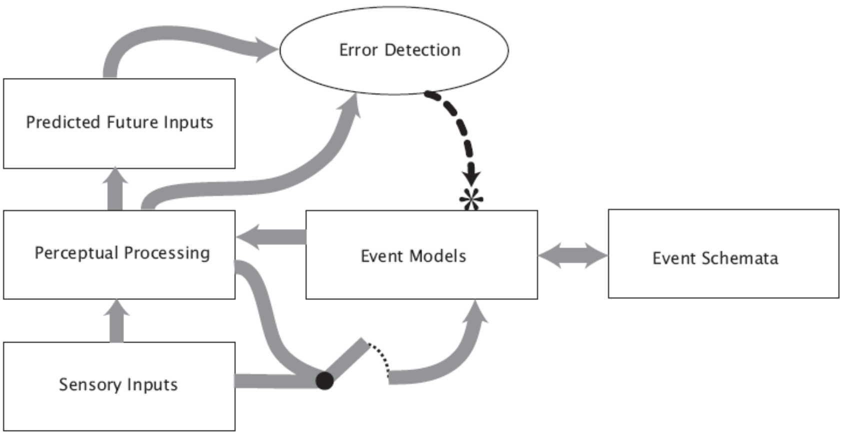


Figure 3.5. Schema of the perceptual processing stream
(Zacks, Speer, Swallow, Braver, and Reynolds, 2007, pages 274)

Figure 3.5 describes the flow of information between processing areas which are indicated by the thin gray colour arrows. Dashed lines signal projections that lead to the resetting of event models. The connection from sensory inputs to event models is gated, such that even models receive sensory input only during the reset phase. Event models also receive input from event schema which contains previously learned information about the sequential structure of activity. Future events, prediction quality are evaluated by an error detection mechanism that compares the perceptual processing stream’s prediction with what actually happens in the world (Zacks, Speer, Swallow, Braver, and Reynolds, 2007).

3.5.2. A typical example of perceptual processing: Newton’s cradle

Newton’s cradle consists of a series of identically sized metal balls suspended to a rack by wires, such that they line up and are in contact when they are in resting position (Figure 3.6).

Newton’s cradle is named in honour of scientist and mathematician Isaac Newton who founded Newton’s Laws. This device is widely known as Newton’s ball or Newton’s pendulum that demonstrates the Laws of Conservation of Energy and Momentum (Gavenda, & Edington, 1997).

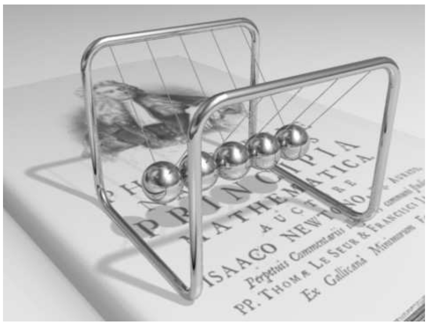


Figure 3.6.Example of “Newton’s Cradle” http://wikipedia.org/wiki/Newton’s_cradle

How the Newton’s cradles work? When the ball in the end position is pulled up from the other balls then released, it swings down and hits the next ball (see in fig.3.6). If two or more balls are pulled up and released, the same number of balls will be propelled forward on the other end, due to the collision. The action will go back and forth until it progressively slowly slows down due friction losses and the elasticity of the balls. The action of Newton’s Cradle is explained and summarized in Figure 3.7.

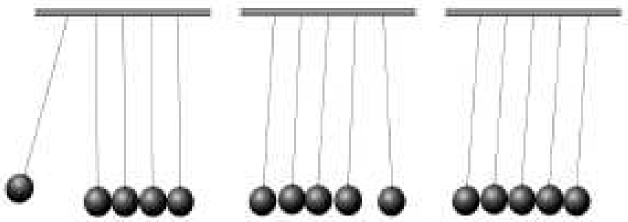


Figure 3.7.Three Phase of Newton’s Cradle dynamics: (1) ‘end bounce’, (2) ‘middle jostle’, and (3) ‘joint swing’. Source: Lowe (2006)

Lowe (2006) tested the Newton’s cradle presentation to know how temporal manipulation of such animated depictions may affect learners perceptual processing in different levels of dynamic information. In the experiment, the learners were divided in two groups, one group was shown a normal speed version and the other was shown a half speed version. They were asked to explain how Newton’s cradle worked and to think while this sequence was showed to them individually on three successive occasions. Verbalisations and eye tracking data were used in that experiment. Result showed that verbalisations from learners in the normal condition tended to be more descriptive than explanatory overall, the contradiction applied to slow condition group. Eye tracking data indicated that in learners in the slow speed group directed more their attention resources to the crucial micro level information within the animation than those in the normal speed condition.

3.6. The Animation Processing Model: APM

Lowe and Boucheix (2008) proposed a model of processing visual information and processes that involved in learning from animated graphics. The first step in building a mental model of a dynamic system is the parsing of the continuous flow of dynamic information into event units. The learner explores the overall animation at local levels. This first step is based on: (1) the ability of the cognitive system to detect spatially the elements constituting a visual contrast in relation to other elements, (2) the ability to detect and encode among a set of moving parts, made salient temporal discontinuities in that they are "dynamic contrast", and (3) the identification of temporal discontinuities and dynamic contrast in a second step allows the hierarchical decomposition of the system in different sequences of events. This decomposition is based on the principles of categorization (Rosch 1978, in Schnotz & Lowe, 2008), and the subject attempts to anticipate the following information (Boucheix & Lowe, 2008).

The animation with their isomorphism and mechanism or phenomenon represented, have a function as "representational"; they allow describing direct and explicit local movements and changes of spatial-temporal in dynamic system (Schnotz & Lowe, 2008). They can facilitate the mental simulation of this system. The influence of perceptual and attentional processes is dominant in the first phase of building a mental model, particularly when subjects have a lack of knowledge in the specific areas presented (Kriz & Hegarty, 2007; Schnots & Lowe, 2008).

A theoretical model describing learners perceptual and conceptual processing of animation was proposed by Lowe & Boucheix (2007, 2008). The model is composed of five hierarchical stages in the Animation Processing Model (APM) as described in below (Figure 3.8).

Animation Processing Model

Top down influence	**Phase 5: Mental model consolidation** Elaborating system function across varied operational requirements *Flexible high quality mental model*
	Phase 4: Functional differentiation Characterization of relational structure in domain-specific terms *Functional episodes*
	Phase 3: Global characterization Connecting to bridge across 'islands of activity' *Domain-general causal chains*
	Phase 2: Regional structure formation Relational processing of local segments into broader structures *Dynamic micro-chunks*
Bottom up influence	**Phase 1: Localized perceptual exploration** Parsing the continuous flux of dynamic information *Individual event units*

Figure 3.8.Animation processing model (Lowe & Boucheix, 2011)

During stage one; learners try to parse the continuous flux of dynamic information into individual events. The APM characterises phase one processing in terms of the parsing of an animation's continuous flux of dynamic information into a set of discrete even units (i.e., entities plus their associated behaviours). This initial parsing is accomplished during perceptual exploration at a very local level. Once individual event units have been identified and internalized, phase two processing can relate the local segments into connected regional structures (i.e., forming dynamic micro-chunks). As indicated in Fig.3.8, the emphasis in phase one and two tends to be on bottom-up processing. Phase three provides more detailed description of specific aspects, such as the sequencing and the duration of dynamic changes. The events that were identified in phase three are interpreted and characterized as functional relations in phase four. Phase five is a mental model consolidation stage, which elaborates the system functions between several causal chains of the representation in order to built a higher quality functional mental model.

Lowe and Boucheix (2008) used as an example of complex animation the upright piano mechanism. This mechanism is composed of three functional stages in mechanism's operational cycle to illustrate aspects of the theoretical model (Figure 3.9).

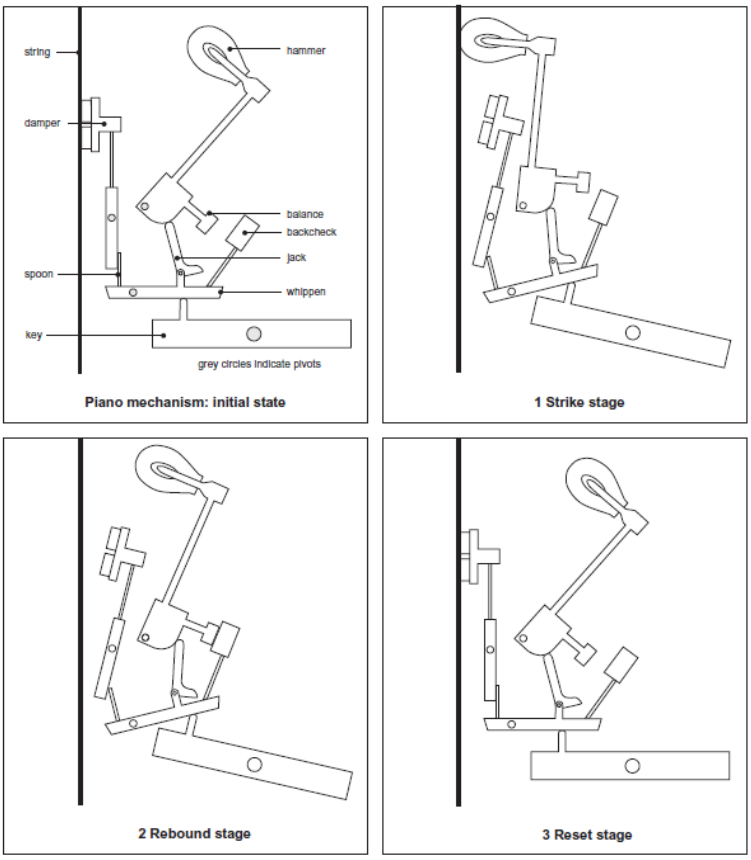


Figure 3.9.The components of an upright piano mechanism and three functional stages in piano mechanisms (Boucheix and Lowe , 2008)

Three functional stages consist of stage 1 (*Strike*), stage 2 (*Rebound*), and stage 3 (*Reset*) (Figure 3.9). Each of the piano's keys operates such a mechanism in order to produce its corresponding musical note. When the key is pressed down to play a note, its opposite end tips the whippen anticlockwise. As a result, the spoon pushes the tail of the damper causing it to move clear of the string. Pivoting of the other end of the whippen in the opposite direction causes the jack to push the hammer forwards. It strikes the string just after the damper has been retracted and produces the required musical note (Stage 1). Having struck the string, the hammer rebounds but its backward movement is arrested when the balance reaches the backcheck. Then, for as long as the key remains depressed, the mechanism stays 'frozen' in position (Stage 2). Once the key is released, the mechanism's components return to their starting places and are reset ready to respond to the next key press (Stage 3) (Lowe and Boucheix, 2008).

CHAPTER 4 *INDIVIDUAL DIFFERENCES*

A lot of previous study found significant effect of spatial abilities and prior knowledge on graphic processing. Visual-spatial abilities could be a requirement for the perception and comprehension of static and dynamic process from external visualization. Hegarty (2005) offered the hypothesis that in learning with dynamic visualisations, spatial ability might play the role of an enhancer: learners with high spatial ability might profit from learning with animations, while learners with low spatial ability might not. However, another plausible hypothesis stated the possibility of compensating effect for low spatial ability: learners with low spatial ability might be supported by dynamic visualizations, because the visualization provides the learners with an external representation of a process that helps them build an adequate mental model; it should be more difficult to build such a model by using static pictures (Hays, 1996). In this short chapter, the concept of spatial ability and prior knowledge are presented. A short review of research on spatial ability and prior knowledge on learning with visualisation is proposed.

4.1 The concept of spatial ability

Spatial ability could be an individual factor that can influence a learner's ability to extract information from dynamic visualization. Mostly, spatial ability correlates with individual abilities in searching the visual field, recognizing forms, shapes, and positions of objects visually perceived, forming mental representations of those forms, shapes, and positions, and manipulating such representations "mentally" (Caroll, 1993).

There are five factors which were identified by Carrol (1993) in the category that referred to abilities in the domain of visual perception: (1) spatial visualization, (2) spatial relations, (3) closure speed, (4) flexibility of closure, and (5) perceptual speed.

Spatial visualization is the ability to mentally manipulate, rotate, or invert objects (Hegarty and Waller, 2005). Frequently, the typical task in spatial visualization is the paper-folding test. The task measures the ability to imagine a result after folding or assembling parts of an object. *Spatial relations* are the ability to estimate changes in the orientation of an object that usually involve simple (two-dimensional) mental rotations items (Lohman, 1979). The card rotations test is a measure of spatial relations that refer to speed rotation.

Closure speed is related to access to spatial representations quickly from long-term memory. The Snowy Picture Test is a test of closure speed in which the task is to identify the object in the picture as quickly as possible.

Flexibility of closure is more closely related to the ability to hold a stimulus in working memory with attempting to identify it from a complex pattern. *The Hidden Figure Test* is a test of flexibility of closure which consists in a task of discrimination of complex figures.

The Perceptual speed was described by French (1951) as speed in comparing figures or finding a known visual pattern. *The Identical Pictures test* is used for test of perceptual speed. Sample items from tests of different spatial abilities are shown in Figure 4.1.

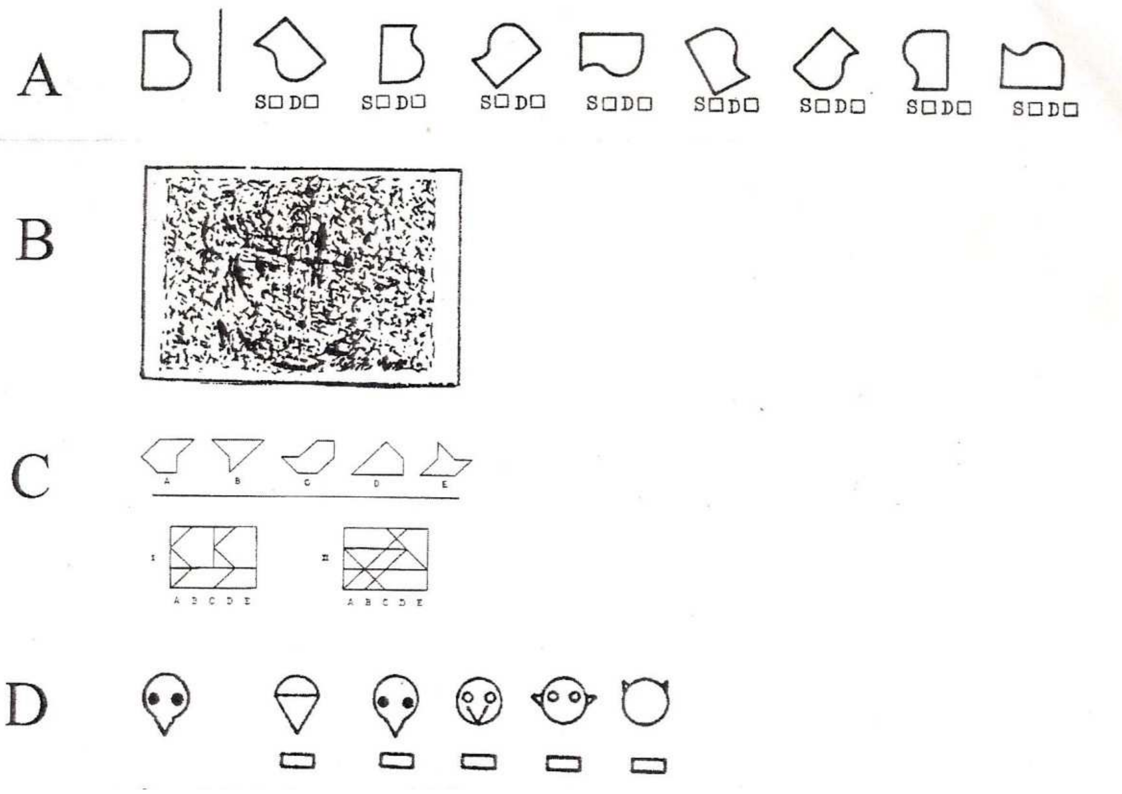


Figure 4.1.Sample items from tests of different spatial abilities factors identified by Carroll (1993), taken from Ekstrom et al., (1976): (A) The Card Rotation test, (B) The Snowy Picture test, (C) The Hidden Figures test, (D) The Identical Pictures test (Hegarty and Waller, 2005, pages130)

Figure 4.1 shows that in The Card Rotation test (A), subject must determine which of the items on the right are rotations of the item on the left of the line (as opposed to its mirror image). In The Snowy Picture test (B), the participant must identify the object in the picture (in this case, an anchor). For The Hidden Figures test (C), subject must determine which of the shapes, labelled A to E, and is contained in each of the complex figures below. The Identical Pictures test (D), subject must identify which of the picture on the right is identical to the one on the left.

Hegarty (2004) explained that animations and static diagrams can be thought of as external visualization (external visual-spatial representations) and spatial ability can be thought of as internal visualization ability. Interaction between perception of external visualisations and internal visualisations processes may determine the quality of spatial mental models.

How does spatial ability influence learning from animations? Bennett (1969) stated that spatial ability has been found to be highly correlated with mechanical ability, which is not surprising given that an accurate mental mode of mechanical systems must include a representation of spatial properties such as shape of components, configurations, and movement. People with high spatial ability are better able to understand mechanical processes even without instruction. There are different option to describe the interactions between spatial abilities and the format of instruction. One possibility is that an animation may act as a cognitive "prosthetic" for those with low spatial abilities. This idea allows predicting that low-spatial individuals would learn relatively more from animations than high-spatial individuals. A second possibility is that spatial abilities may be a prerequisite for accurate perception and comprehension of external visualization. A third possibility is that external visualization augments spatial abilities. In the option it is expect that external visualization would support both the high-spatial and the low-spatial individuals (Hegarty & Kriz, 2008).

The meta-analysis realized by Höffler (2010) was focused on the role of spatial ability in learning from animated visualizations. There were 19 studies (published between 1999 and 2009) which were included in the meta-analysis. Result showed that spatial ability plays an important role in learning from visualizations (mean effect size of $r = 0.34$), but this effect is moderated by two compensating factors: learners with low spatial ability can be significantly supported by a dynamic visualization as well as a 3D-visualization. It confirmed that spatial ability is a factor which should be considered when designing visualization experiments. Learner with low spatial ability can be supported by some design modifications of external visualizations.

The recommendation to use dynamic visualization for learners with low spatial ability may look controversial at first glance, however this could be a good reason coming from many different findings concerning static pictures versus animations in the past (Höffler and Leutner, 2007; Mayer et al. 2005; Tversky et al. 2002). Learners with high spatial ability could learn better when they face with non-transient static pictures, which give them the opportunity to construct a mental model as a result of their highly developed spatial ability. Learners with low spatial ability could learn better from animations; maybe because animations will provide them with a ready-made dynamic model of the process presented. Dynamic visualizations have a compensating effect for learners with low spatial ability.

4.2. Prior Knowledge

Prior knowledge is a more classical factor found to be influenced in many learning areas. Chi (2000) classified four types of prior knowledge that learners may possess: (1) domain-specific knowledge, (2) domain-relevant knowledge, (3) misconceptions, and (4) domain-general world knowledge. Considering effects of prior knowledge on learning, it is important to differentiate domain-specific prior knowledge from domain-general knowledge, and not simply distinguish “high” or “low” levels of prior knowledge. Knowledge acquired from physics classes may have different effects on learning in comparison with knowledge acquired from practical experience interacting with machines (Hegarty & Kriz , 2008).

For an example, Hegarty & Kriz (2008) found that animations compared to static presentation were relatively more effective for low-knowledge individuals, because animations show the motion in a mechanical system explicitly, and do not depend on the learner’s ability to infer motion from the static diagrams.

However, alternatively there are some reasons to expect that comprehension of animations may be improved by prior domain knowledge. Constructivist theories of comprehension assume that learning involves the integration of new information into existing knowledge structures (Chi et al., 1994; Kintsch, 1988; Schnotz, 2002). The result of the integration process depends on the quantity, specificity, and accuracy of the existing knowledge besides on how the new information is presented.

Miyake (1986) has suggested that building a mental model from a mechanical system presentation is iterative process of comprehension of how the system works. Revising a mental model can only occur when a conflict between the external representation and the internal representation is perceived (Chi, 2000). A learner with a high level of prior domain knowledge should be better able to evaluate the "gaps" in conflicts between internal and external representations. To the contrary, low domain knowledge learners may be overconfident in judging their comprehension of mechanical phenomena (Rozenblit and Keil, 2002).

Several researchers have suggested cognitive design principles which overcome some problems of low knowledge of individuals engaged in learning from animations. Additional visual features such as cues, arrows, and language are expected to help low-knowledge individuals by directing attention to relevant or specific parts or events of the process (Faraday & Sutcliffe, 1997; Hegarty & Just, 1993; Mayer, 2001).

CHAPTER 5 *WHAT DO WE KNOW ABOUT ON LINE PROCESSING OF ANIMATION: the use of eye tracking technique*

The goal of this chapter is to analyse how eye tracking could be used to investigate visual information processing. This chapter is focused on the use of eye tracking techniques and methods to study cognitive processes during visual learning task, especially during learning complex animation.

Why does research should use an eye tracking methodology? Because these techniques provide an on line protocol of the encoding and attentional processes carried out during the learning time such as: what is attended first and next, and for how long does people spend more time for attention, how much switching of attention is done between different elements of the learning materials, what elements are linked together during attentional switching (Hyönä, 2010). Mayer (2010) explained that the research using eye tracking offers a unique contribution particularly concerning the subject's perceptual processing during learning.

5.1. Theory of Eye tracking

Eye tracking is increasingly being used to provide insight on a variety of tasks involving visual perception. It has been used for nearly one hundred years in psychology. The main area which investigated eye movement was focused on recording eye movements in reading research (Rayner, 1998). Going along with technological advances in the recent years, eye tracking has become a promising tool in order to answer questions about multimedia learning. Eye tracking is a technique used to determine where a person is looking at. It can give information about how learners process the visual information and how they inspect the learning material. An eye tracking system records how an individual's eye movements are distributed so the researcher knows where both eyes are looking at and what information is being acquired through vision. The concepts underlying eye tracking are simple: track the movements of the user's eyes and note what the pupils are doing while the user is looking at a particular feature.

Studies about eye movements are numerous and popular in different research areas, from reading to driving and many other domain. Literature about eye tracking applications can be overviewed and founded in Duchowsky (2007).

Why is eye tracking important? If we can track someone's eye movements, we can follow along the path of attention deployed by the observer. This may give some insight into what the observer found interesting, what drew their attention, and perhaps even provide a clue as to how that person perceived whatever scene she or he was viewing. Eye tracking will play an important role in understanding how people interact with complex displays (Duchowski, 2007).

There are over ten different types of eye movements, of which the most important ones are saccades and fixations. When the eyes focus on a point, it is called a fixation and the movements between these fixations are called saccades. When the eyes fixates, the stops vary from about 100-600 milliseconds and during this stop the brain starts to process the visual information received from the eyes. The length of a fixation is usually an indication of information processing or cognitive activities as this is when the brain interprets the visual information from the eyes (Poole and Ball's, 2005). Some of the most used fixation and saccade derived metrics are described on Table 5.1 and Table 5.2, based on Poole and Ball's (2005).

Table 5.1. Some of the most used fixation-derived metrics. Source: Poole & Ball, 2005

Eye-Movement Metric	**What it Measures**
Number of fixations overall	More overall fixations indicate less efficient search.
Fixations per area of interest	More fixations on a particular area indicate that it is more noticeable, or more important, to the viewer than other areas.
Fixation duration	A longer fixation duration indicates difficulty in extracting information, or it means that the object is more engaging in some way.
Gaze	Gaze is usually the sum of all fixation durations within a prescribed area. This is used to compare attention distributed between targets. It can also be used as a measure of anticipation in situation awareness if longer gazes fall on an area of interest before a possible event occurring.
Percentage of participants fixating an area of interest	If a low proportion of participants fixating an area that is important to the task, it may need to be highlighted or moved.

Table 5.2. Some of the most used saccade-derived metrics. Source: Poole & Ball, 2005

Eye-Movement Metric	**What it Measures**
Number of saccades	More saccades indicate more searching.
Saccade amplitude	Larger saccades indicate more meaningful cues, as attention is drawn from a distance.
Regressive saccades (regression)	Regression indicate the presence of less meaningful cues.

Area of Interest (AOI) is also a common method used in eye tracking. In our research, define areas of interest over certain parts of a display or interface under evaluation and analyse only the eye movements that fall in that areas. AOI allow researchers to make conclusions about how often a specific object was looked at and in what order the different objects were inspected.

Another visualization method for eye movement data are scan path diagrams and heat maps. Scan path is an eye tracking metric that usually consists of a complete sequence of fixation and interconnectivity saccades. They use dots to indicate the fixation duration and lines to indicate length and direction of saccades. Heat maps are colour coded representations of aggregated eye tracking data. It consists of a transparent background with highlighted areas where the test subjects have been looking. Scan path and heat map do not need deeper understanding of eye tracking metrics but when used in quantitative research, they are difficult to interpret.

5.2. Eye tracking in research on multimedia learning

The use of eye tracking as a tool in the study of multimedia learning has been developed to provide empirical evidence about aspects of theories in multimedia learning.

Eye tracking can provide insight into how people understand and learn about complex systems in multimedia learning. Most of examples come from the domain physical systems include man-made mechanical systems and biological systems (e.g. pumping blood in the case of the human heart, weather maps and piano animation for how mechanical devices work) (De Koning, Tabbers, Rikers, Paas, 2010; Canham and Hegarty,2010; and Boucheix and Lowe, 2010).

Ozcelik, Karakus, Kursun, and Cagiltay (2009) investigated how color coding affects multimedia learning by utilizing eye movement data. Data demonstrated that average fixation duration was longer when the material was coded by colors. Longer average fixation duration is noticed as an indication of deeper processing (Rayner, 1998), because fixation duration is thought to be determined by current cognitive processes. When it was easy for learners to find relevant information and integrate verbal and pictorial information with each other, learners could more easily take part in deeper processing required for meaningful learning (Mayer, 2003; Mayer & Chandler, 2001).

Using eye tracking data the colour coding assists learners to find related information in text and illustration and to pay attention to critical information for meaningful learning. Based on their results, authors suggested that designers should use the same colour to associate elements in related text and illustrations. To enhance learning the attention of learners should be directed to essential elements in the instructional material, and unnecessary search for related information should be minimized.

Boucheix & Lowe (2010) used eye tracking to investigate a novel cueing approach for directing learner attention to low salience, high relevance aspects of a complex animation. They used spreading colour cues that highlighted the relevant features of an animation depicting how a piano mechanism works was compared with a piano mechanism animation containing arrow cues or no cues. In their study learners spent more time looking at relevant areas of an animation when relevant features are highlighted both spatially and temporally. Examination of eye movements during learning animation confirmed that cueing vs. no cueing directed learners' attention within animation. Eye movement recording indicated that cueing effectively guided attention to the signalled region.

De Koning, Tabbers, Rikers, and Paas (2010) also used eye movements to examine how visual attentional resources are located when learning from a complex animation of the human cardiovascular system. The difference in fixation patterns between the cued and non-cued conditions is taken as proof that cueing guides learners attention to specific regions in an instructional animation.

Previous studies have shown that signaling enhances multimedia learning. However, there is not enough evidence showing why signaling enhances multimedia learning. Therefore, Ozcelik, Ari, and Cagiltay (2010) continued to investigate in depth, the causes of signaling effect by employing time-locked and additional analysis of fixations. Several studies with eye tracking most commonly used total fixation time on signalled information as a measure for perceptual processes during learning (Boucheix & Lowe, 2010; De Koning, et al. 2010). In addition to global eye-tracking measures, Hyönä (2010) proposed that time-locked analyses can complement the global picture derived from total fixation time.

Ozcelik, Ari, and Cagiltay (2010) used material multimedia package included a labelled illustration of a turbofan jet engine (Figure 5.1) and a narration by a female voice, by the speaker of a computer explaining how turbofan jet engines work.

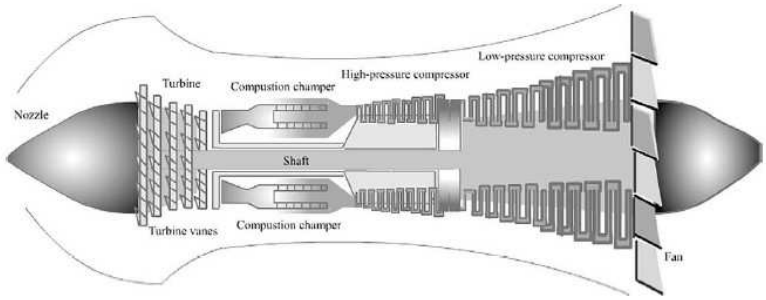


Figure 5.1. The non signal format of turbofan jet engine (Ozcelik, Ari, and Cagiltay, 2010)

The signalled format was identical to the non signalled one with one exception. In the signalled format, each corresponding terminological label was presented in a red colour during the narration of the sentence in which the item was mentioned. The colour of the label was changed to its original colour (black) when the narration ending. Results explained that the signalled group outperformed the non signalled group on transfer and matching tests. However, both groups performed similarly in the retention test, pointing out that the effect of signaling was on deeper processing.

In this study, the eye movement data helped to explore the source of the signaling effect. Two reasons explained by Ozcelik, Ari, and Cagiltay (2010) how the signalling effect enhances meaningful learning: First, there is evidence of a higher number of fixations and longer total fixation time on relevant information including both labels and related parts of the illustration. It means that learners allocated more attention to relevant information when they were guided by signals, which is consistent with previous eye tracking studies (Boucheix & Lowe, 2010; De Koning et al, 2010). Second, the learners spend less time finding relevant labels, because signaling facilitated the efficiency and the effectiveness of visual search to find necessary information.

Study about the effect of knowledge in understanding a complex graphical display also using eye fixations analysis was proposed by Canham & Hegarty (2010). Weather maps were used as complex graphical displays. The learners were taught the relevant meteorological principles, and then they made inferences from weather maps.

Results from the measures of eye fixation time and cognitive performance showed learners with high prior knowledge performed better on eye fixation time scores and cognitive performance scores than did learners with low prior knowledge. Mayer (2010) stated that there is proof of a relation between increases in perceptual processing of relevant portions of the graphic and enhancement in measures of cognitive performance on an intellectually demanding task.

Weigand, Kohnert, and Glowalla, (2009) investigated visual attention distribution in learning from text and pictures using eye tracking. The goal of their experiment was to examine the effect of instruction pace on viewing behaviour: how do the learners manage to split attention in system-paced and in self-paced instruction, and how does text modality affect the viewing behaviour in system-paced instruction. The learning material was based on an animation used by Mayer and Moreno (1998), and consisted of 16-steps multimedia instruction on the formation of lightning programmed in Flash 4.0. Visualizations were designed to explain the dynamics of the major events, for example describing the motion of cool air that becomes heated, drops of water and ice crystals moving up and down within the cloud, and positive charges moving up to the cloud producing a flash light. Each step was visualized in a single scene and text previewed in German. Sample screenshots for the written text presentation are shown in Figure 5.2.

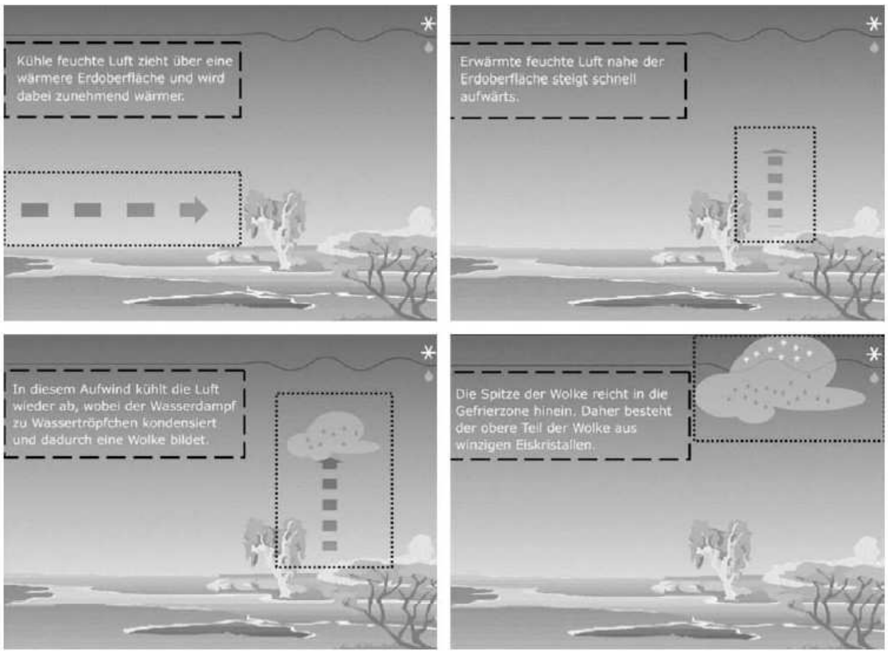


Figure 5.2. Areas of interest (AOI) for visualization (dotted rectangles) and written test (dashed rectangles) for selected frames of the multimedia instruction (scenes 1-4) (F.Schmidt-Weigand et al., 2009)

Results showed that learners spent more time studying the visualization with spoken text than those with written text. Learners consistently started reading text before changing between text and visualization, and they spent more time reading the text than inspecting the visualizations. Learners in system-paced instruction used additional presentation time in favour of visualizations, while learners in self-paced instruction used additional presentation time for reading. Furthermore, learners in system-paced instruction used additional presentation time to change their visual attention between text and visualization more often, while learners in self-paced instruction did not. Related to learning outcomes (retention, transfer, and visual memory), is an effect of the text modality only for visual memory. In conclusion, they stated that the distribution of visual attention in multimedia learning is mostly directed by text.

Mayer (2010) examined the six studies using eye tracking as a tool to study and enhance multimedia learning. Research using eye tracking offers a unique path to testing aspects of theories of multimedia learning. Table 5.3 lists some of main features of the experimental comparisons reported by Mayer (2010).

Table 5.3. Some of the main features of the six-eye tracking studies (Mayer, 2010)

Study	Content	Medium	Main Independent variable	Eye fixation measure	Main learning outcome measures
Boucheix and Lowe, (2010)	Piano	Animation only	Signaling: spreading colors vs. Arrows vs. No arrows	Time looking at relevant areas	Comprehension score (including configuration, local kinematics, and mental model score)
De Koning et al., (2010)	Heart	Animation only	Signaling: shading cues vs. No cues	Time looking at relevant areas	Comprehension and transfer scores; number of explanatory statements in verbal protocols
Jarodzka et al., (2010)	Swimming fish	Video only	Expertise: professors vs. students	Time looking at relevant areas	Accuracy in describing locomotion
Canham and Hegarty (2010)	Weather maps	Static map only	Prior knowledge: before vs. After instruction	Time looking at relevant areas	Accuracy in verifying wind direction
Schmidt-Weighand et al., (2010)	Lightning	Animation and words	Modality: animation with narration vs. Animation with on screen text	Time looking at relevant areas	Comprehension, transfer, visual memory scores
Meyer et al., (2010)	Four-stroke engine	Animation only	Presentation rate: fast vs. slow	Time looking at relevant areas	Comprehension score

In table 5.3, all of the studies involve graphics, including animation, video, and static illustrations, but only one involves accompanying words (Schmid-Weigand et al., 2010). Total fixation time on relevant areas of a graphic is used in eye fixation measure; function as dependent measures and dependent measures involving learning outcome or cognitive performance such as transfer test score.

The authors of the six-studies have been able to show that signalling effect, prior knowledge effect, and the modality effect can be extended to measures of eye-fixation time (Mayer, 2010).

5.3. Difficulties and constraints of eye tracking method

Besides many benefits of eye tracking, there are some constraints and possible difficulties of the eye tracking system. Eye trackers are quite sensitive instruments and can have difficulty tracking participants who have eye-wear that interrupts the normal path of a reflection, such as hard contact lenses, glasses with super-condensed lenses. Some people cannot be eye tracked for physiological reasons: people with very large pupils or “lazy eye” or their eyelid obscuring part of the pupil and making the eye difficult to track. Eye tracking technology needs to be improved to increase the validity and reliability of the recorded data. The robustness and accuracy of data capture need to be increased, so that the point of regard and measurement stays accurate without the need for frequent re-calibration (Poole and Ball, 2005).

Problems can occur during the eye-tracking process when a person’s eyes may dry out during an experiment and become difficult to track. Sometimes a participant can be eye tracked one day and not the next, or half way through a test the eye track degrades to the point that the data collected cannot be used. (Namahn, 2001).

Some techniques of eye tracking can be uncomfortable for the user because of the equipment such as special contact lenses, electrodes, chin rests, head mounted or other components that must be physically attached to them.

5.4. Dynamic Ocnoscere

Ocnoscere is a software program created by Casallas (2008) which was implemented in our research (Putri, Boucheix, & Cassalas, 2009) in LEAD. The name of *ocnoscere* comes from two Latin words: knowing “cognoscere” or “noscere” and eye “oculus” or “oclus”.

Ocnoscere means “knowing through the eyes”. Following a simplified version of the Latin phonology, using the International Phonetic Alphabet (IPA), this word should be pronounced [oknoʃere]. *Ocnoscere* is a new tool to analyze dynamic areas of interest (AOIs), because previously research was only used to analyze static areas of interest. Ocnoscere allows the analysis of two and three dimensional animations using eye-tracking. It was initially developed to work with the Tobii 1750 screen-mounted video eye tracker, although it can be used with other Tobii eye trackers which the support of the Tobii software development toolkit (SDK).

Ocnoscere uses an area-based algorithm to identify gaze points within predefined areas of interest (AOI) (Salvucci 2000); these AOI correspond to the different parts of the 3D model loaded on the program. Using the user’s raw gaze input from the eye tracker, the algorithm calculates the first intersection between the ray originating in the middle of the user’s eyes and the loaded 3D model, in a basic ray-casting or picking operation (Figure 5.3).

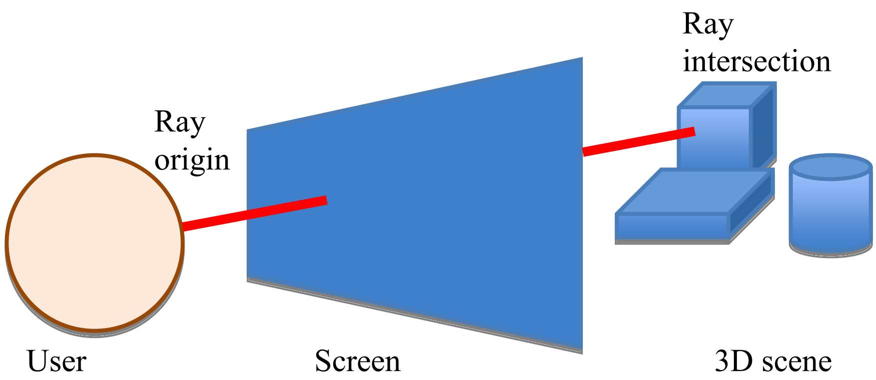


Figure 5.3.Ray-casting algorithm in *Ocnoscere*

Since the AOIs are proper to the 3D model, *Ocnoscere* makes the eye-tracking analysis online, i.e. during the recording, which allows analyzing not only static, but also animated 3D models. This approach is radically different from eye-tracker-furnished eye tracking analysis software in which AOIs can only be defined in a static way, thus requiring a post hoc analysis.

The program allows reviewing the results of one or many previous eye-tracking recordings on the model. For this, the number of raw gaze data per AOI is shown in a numeric way, and the 3D model is re-colored using the HSV (Hue, Saturation and Value or lightness) color scale, leaving S and V constant and varying only over H, according to the normalized value of raw gaze data. This reviewing pane can be seen in Figure 5.4.

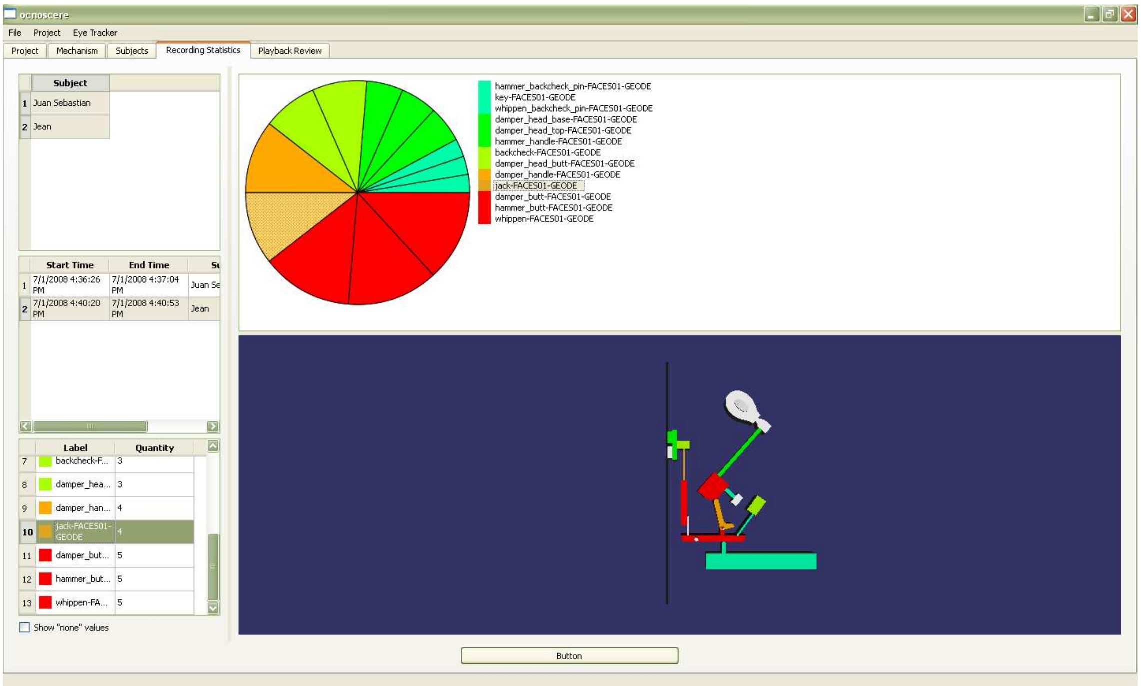


Figure 5.4. *Ocnoscere* review pane showing raw gaze data. (a) On the upper-left corner, the subjects, (b) on the middle-left the user's recordings (c) on the bottom-left the AOIs with their associated values, both numerically and in the HSV scale. (d) On the upper pane, a pie-chart summarizing the results and (e) on the bottom pane, the model re-colored according to the raw gaze data.

Additionally, each recording can be revisited; showing the animation and camera movement together with the user's scan path and the re-colored model as shown in the review pane (see Figure 5.5). All of the results can be exported to comma-separated text-files, to allow for further data analysis and exploitation in separate programs. We used this software in our studies (experiment 4) as a tool for eye tracking analysis on dynamic areas of interest (AOIs).

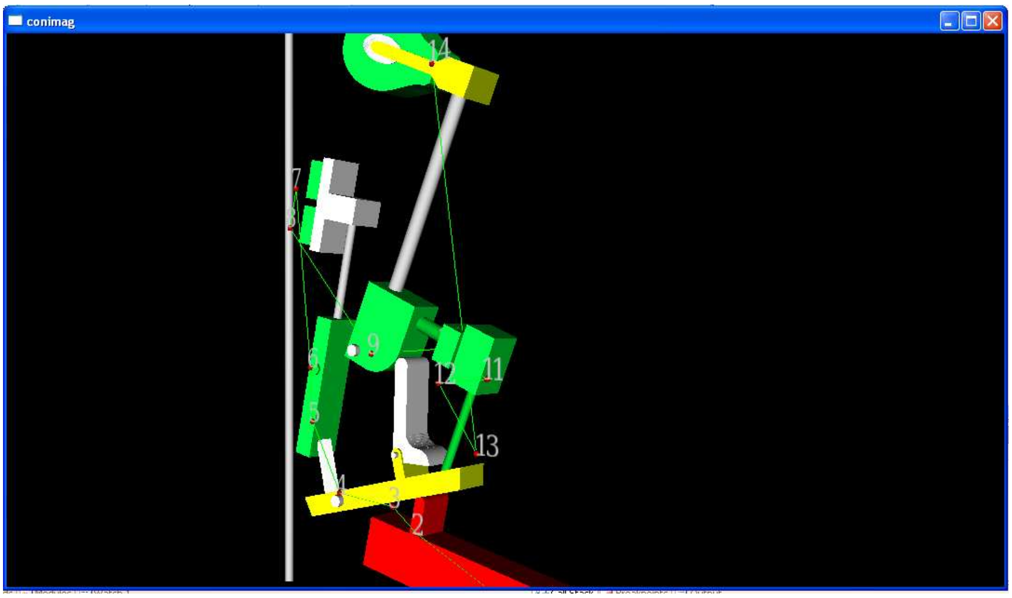


Figure 5.5. Scan path review for a single recording in *Ocnoscere*

PART

2

EXPERIMENTAL

CHAPTER 6. Research Goals

6.1. Introduction

Numerous researches about animation were mainly focused on the effect of external format of the animation or static pictures on comprehension performance. The goal of these researches was not always focused on how mental model are built.
Previous studies used different types of simplified representations of the dynamic referent content, such as series of static pictures or animation segments to depict the key steps of the animation (Kriz & Hegarty, 2007, Arguel & Jamet, 2008, Rebetz et al, 2009, and Betrancourt & Morand, 2010).

Our studies had four main questions related to four specific goals. The first goal consisted in testing new segmentation techniques using dual format presentation (Experiment 1); the second goal, was focused on the elaboration and the test of new cueing techniques applied on dual format presentations (Experiment 2 and 3). The third goal was to compare 2D/3D format presentation (Experiment 4). Finally, the fourth goal was to investigate the possible role of executive function in animation comprehension (in Experiment 2 and 3).

Usually, series of static pictures are not always better than animation (see theoretical part). In experiment 1, two integrated dual format presentations were compared to single format presentations. We tested four presentation formats of an animation of an upright piano mechanical system: (i) a series of static pictures followed by an animation, (ii) an animation followed by a series of static pictures, (iii) an animation followed by an animation, and (iv) a series of static pictures followed by a series of static pictures. It was predicted that when a participant is studying a series of static pictures (showing the key steps of the dynamic process) followed by an animation, it is quite possible that the inferences built from the key steps of the series of static pictures would be checked by the learner during the cognitive processing of the following animation. This activity could enhance high quality mental model building.

Previous research used mainly two forms of cueing: Static cueing of a single component (colours, arrows, etc) and dynamic cueing of animated processes (spreading colours cues:

Boucheix and Lowe, 2010). In Experiment 2 and 3, new cueing techniques were applied by using spatio-temporal colored “tokens” with series of static pictures followed by an animation. This technique was supposed to guide the learners’ cognitive processing to establish a link between the key steps of the process and the animation. Similarly to the first experiment, we analyzed four experimental conditions of the same of piano mechanism system and a control condition. Eye movement of participants were recorded during the experiments 2 and 3, in order to follow more precisely on line cognitive processing which could give new insights on how mental models are progressively built.

In experiment 4 was created another type of presentation using 2D and 3D animations as alternative cueing technique. Two D and 3D animations were compared. Because 3 D presentations could augment the perceptibly profile of relevant components better comprehension was expected in 3D compared to 2D.

In this experiment, “Ocnoscere” was used. “Ocnoscere” is new software for eye tracking data processing, especially designed for the construction and analysis of dynamic area of interest. This is the first experiment which used such software for the analysis of dynamic area of interest.

In three of experiments (2, 3, 4), spatial abilities were controlled and measured. The effect of spatial ability and animation comprehension was analyzed. The effect of spatial abilities as a moderator in animation processing is now well-known (Höffler & Leutner, 2010). However, the cognitive treatment of an animation could also require: (a) a high speed processing of transient information, (b) the continuous refreshment of current representation of the dynamic process, and (c) the inhibition of no relevant information from the presentation. These abilities of working memory are called executive functions of the brain. Executive function could play a role in animation processing performance. In our research, we proposed to use specific measures of information processing, speed processing, refreshment and inhibition, to investigate the possible effect of executive function in working memory on the understanding of complex animation.

Four tests were used. The Stroop test (Stroop, 1935), was designed to measure the capacity of inhibition. The N-Back test (Kirchner, 1958), was used as test of general capacity for maintaining and updating information in working memory. The Plus-Minus test (Spector &

Biederman, 1976) was used to investigate attention switching from one task to another and the measure of cognitive flexibility. Finally, the XO test (Salthouse, 1990), was used to measure working memory speed processing.

6.2. Overview of the experiments

In table 6.1 a summary of the goal of the four experiments is presented.

Table 6.1. A summary of studies the understanding of learning with animation

No	Format Presentation	Treatment	Prediction
1.	Dual format presentation: Chaining two types of formats in order to foster inferences in building high quality mental models.	Two presentation formats: a series of static pictures and an animation combined in different positions.	An integrated static sequential presentation followed by an animation presentation should be better for developing a dynamic mental model than single formats presentation. **(Experiment 1)**
2.	Signaling relevant information with spatio-temporal colored "tokens" of dual format presentations.	Cueing spatial temporal "tokens" of series of static pictures with a cued animation and control condition.	The adding of cues in format presentations, a series of static pictures and an animation, affect positively animation comprehension. Strategies of dual format presentation with cueing would result in better comprehension than strategies of cued single format presentations. **(Experiment 2)**
3.	Signaling relevant information with colored "tokens" of dual format presentations.	Cueing spatial temporal "tokens" of series of static pictures with an un-cued animation.	Cueing, with visual tokens, only the series of static pictures and not the animation could foster inference activities and improve comprehension performance. **(Experiment 3)**
4.	Comparison of 2D/3D presentation of an animation with an eye tracking analysis	Comparing two types of perceptibility profile with and without depth.	Three D animation, compared two 2D animations, should improve direction of attention on more relevant and less salient components of the animation. **(Experiment 4)**

CHAPTER 7

EXPERIMENT 1. Chaining two types of formats in order to foster inferences for building high quality mental models.

Introduction

In psychology, the term “mental models” is sometimes used to refer to mental representations or mental simulation. Mental models can be constructed from perception, imagination, or the comprehension of discourse (Johnson-Laird, 1983). Mental models represent entities and person, events and processes, and the operations of complex systems.

In complex animations learners are challenged to extract relevant information from a visual display, select corresponding parts of information, and integrate all of elements into a coherent representation (Mayer and Moreno, 2003). Animations are supposed to be superior to static graphics, especially when learning concerns a chain of events in dynamic systems. Animations do not only depict object, they also provide information concerning object changes and their position over time (Rieber, 1990). However, learners often fail to process animations effectively (Tversky, Morrisson, and Bétrancourt, 2002), resulting in no advantage compared with static visualization (Hoffler and Leutner, 2007) and learning from animations often fails because complex perceptual and cognitive processing overwhelms the learner’s limited processing capacities (Lowe, 2009).

Segmentation would help the analysis of events and the construction of an accurate mental model. The segmentation techniques, which consist in using, instead of the animation, a series of static pictures showing the key step of the dynamic process, could help the mental elaboration of the processes because each key step from the static presentation of a the series of pictures could be more easily manipulated in working memory. In previous research, segmented presentations are mainly used alone or at the same time as the animation (Bétrancourt, Dillenbourg, Clavien, 2008).

The study of a series of static pictures, showing the key step of an a animation, followed by an animation, should improve the apprehension process of the animation (Tversky, 2002).

The series of static pictures could be mentally processed by the learner as simplified models of the dynamic phenomena, more easily “handable” than an animation in working memory. Thus a series of static simplified representations of the key steps, of a dynamic process followed by animation could lead to better elaboration of a high quality mental model which could facilitate learning.

Research Goals

The purpose of this experiment is to examine whether the building of mental models from animation could be improved by using external simplified model representations of the dynamic reference content. We used an alternative approach to examine the effect of the presentation of simplified external representations, on learning from a complex animation.

We manipulated the presence and the location of a series of static pictures and an animation. A complex animation of an upright piano mechanism was used without text (Figure 7.1).

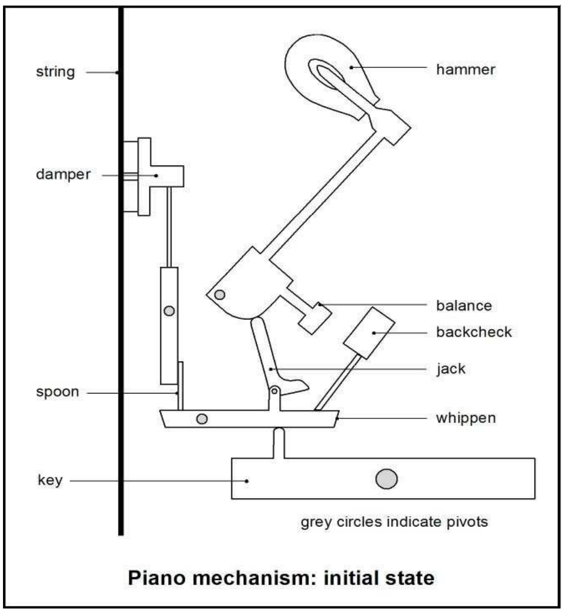
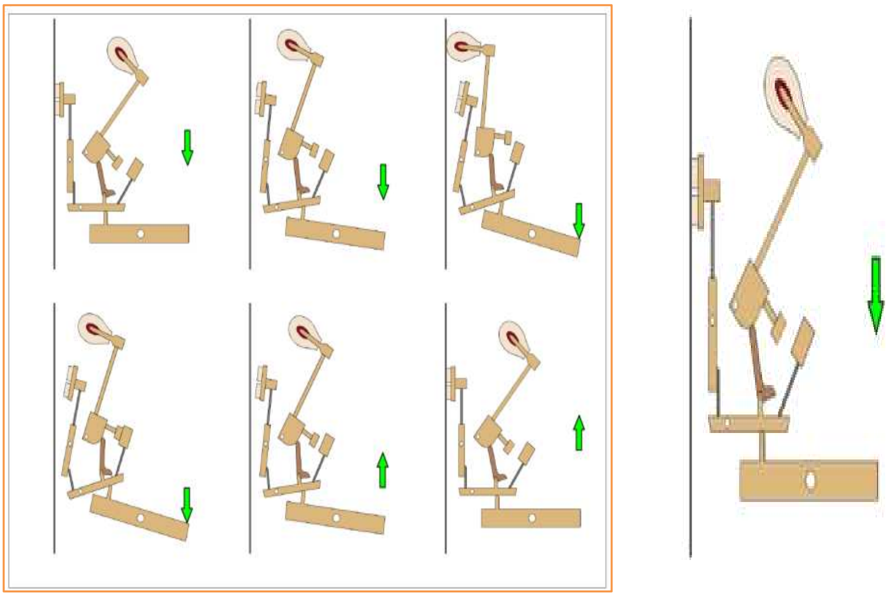


Figure 7.1. An upright piano mechanism (Boucheix and Lowe, 2008)

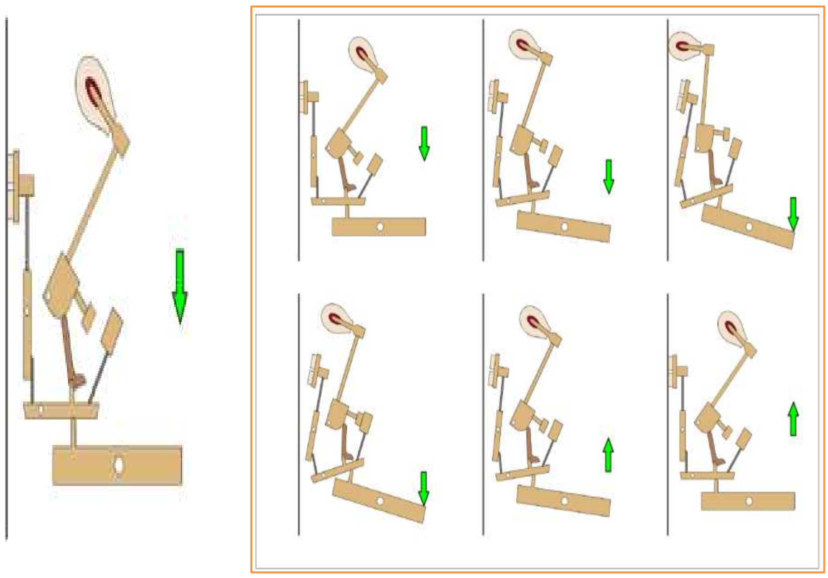
Four types of presentation were compared. In each presentation, the information was delivered in two learning stages (1 minute and 45 seconds for each stage, which means a total learning time of the dynamic process of 3’, 30”). For the first version (Figure 7.2), a series of 6 static pictures (St) depicting the key steps of the mechanism were simultaneously presented during the learning stage 1 and then followed by the presentation of Animation (**An**) during the learning stage2: condition **St+An**.

In the second version (Figure 7.3), the animation was presented first and followed by the 6 series of static pictures: condition **An+St.** The two other versions were controlled modalities. In the **St+St** condition, the same series of static pictures were presented in the learning stage one as well as in the learning stage two, (Figure 7.4). In the **An+An** condition, the animation was presented in the learning stage one as well as in the learning stage two *(*Figure 7.5).



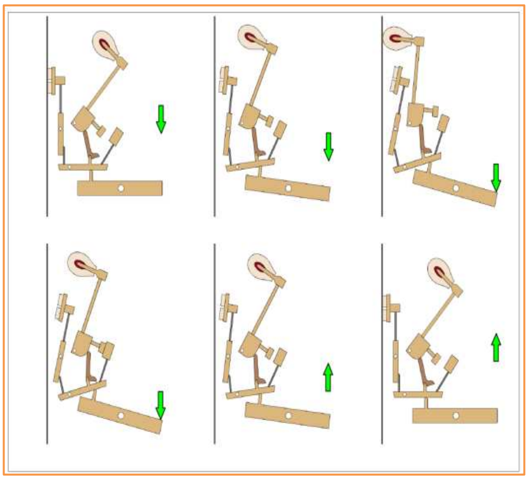
1 minute 45second 1 minute 45 second

Figure 7.2. Series of static pictures and animation (St + An)

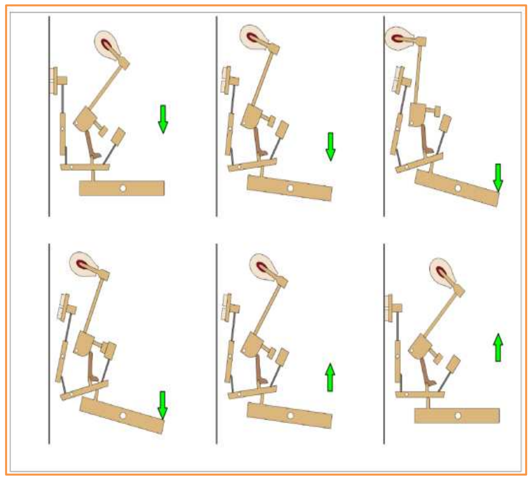


1 minute 45second 1 minute 45 second

Figure 7.3. Animation and Series of static pictures (An + St)

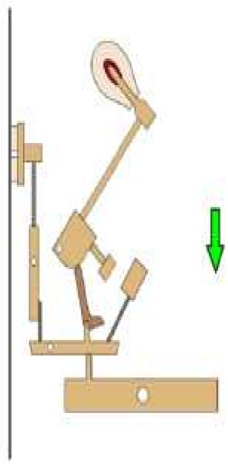


1 minute 45 second

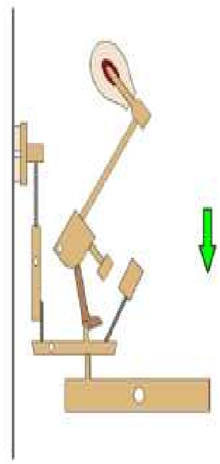


1 minute 45 second

Figure 7.4. Series of static pictures and Series of static pictures (St + St)



1 minute 45 second



1 minute 45 second

Figure 7.5 Animation and Animation (An + An)

Hypothesis

It was predicted higher quality mental models and higher comprehension performances in the **St+An** condition, and to a less extent in the **An+St** condition, than in the single format conditions, **St+St** and **An+An**. It was assumed that the learners would construct a mental model of the mechanism operation by making inferences between the static key steps presented and by using active comparison and mental manipulation of each picture of the series. With the animation, the learners would check the mental model. The animation could give a feedback to the learners for adjusting their representation previously built in the first processing stage of the series of static pictures. For **An+St** condition, the learners could apply the first model built from the animation and make inferences between the static pictures in the second learning stage. For the single format conditions (St+St) and (An+An), there are no such opportunities to 'check' or examine the mental model built in the first learning as an alternative representation.

We predicted that dual format presentations, that is to say, an integrated static sequential presentation with an animated presentation should be better for developing a dynamic mental model than single format presentation.

Method

Participants

Participants in experiment were 82 undergraduate students. They were all between 19 years - 25 years old. Participants were randomly assigned to 4 groups according to four experimental conditions (two dual format and two single format conditions). The design of the experiment is presented on the table 7.1.

Table 7.1. Design of experiment 1

SECOND PRESENTATION	FIRST PRESENTATION	
	Static (S)	Animation (A)
Static (S)	Group SS (N=21)	Group SA (N=22)
Animation (A)	Group AS (N=20)	Group AA (N=19)

Material

A no controllable animation of the upright piano mechanism was used with the four versions presented Figures 7.2. - 7.5.

Procedure

Each participant had to study how the piano mechanism worked for 3 minutes 30 second with 1 minute and 45 second for each animation or series of static picture. This time constraint was chosen in related with previous studies of piano mechanism (Boucheix and Lowe, 2010). For an example: in St+An condition, after one minute 45 second, a series of static pictures is finished previewed then continued by animation previewed. The piano mechanism was presented with pop-up system, when the subject places the cursor on parts of the mechanism, the name of part piano will appear. After viewing the series of pictures and animation, participants were given two post tests.

Comprehension test

Comprehension was measured using two post test. Post-test 1 was the local movement test, which was a direct measure of the recall of the movement of each component.

A novel way of testing kinematic level learning was devised for this study. The *cross movement task* was designed to provide a more direct measure of such learning than is possible using standard approaches such as verbal questions. It required participants to physically produce the movements of the piano mechanism's components and was a computer-based adaptation of the manipulation technique devised by Lowe & Boucheix, 2008, 2011. At the end of the learning session, a static picture of the piano mechanism in its initial state was displayed on the computer screen. As shown in Figure 7.7, a red cross was positioned on a part of a component. The learner was told to *"use the mouse in order to move the cross to the correct final position it occupies when the key is pressed and released*". This required them to demonstrate the movements for stages one (striking), two (rebound), and three (reset). In total, each participant performed the cross movement task thirty times with the cross being on a different position on every occasion. The order in which the crosses were presented (within each stage of the piano mechanism) was randomized across participants. Four examples of these positions are presented in figure 7. 6 and 7.7. For each of the piano mechanism's components, several different cross positions were involved. One cross at a time was displayed on a static picture of the piano mechanism. This comprehension task was assumed to require a mental simulation of the dynamics of each part of the piano mechanism in order to infer movements of the parts shown in the provided static picture. For each position of the cross, the entire movement of the mouse made by the learner was registered by the computer in real time. The angular direction of the movement and its amplitude were subsequently calculated and compared to the actual motion of the component as depicted in the animation in order to determine kinematic precision. The kinematic score for each cross position was based on angular direction (1 point) and amplitude (1 point) of the movement (see Figure 7.7). In this scoring system, the number of points credited to the learner is directly related to the precision of the direction and distance of the movement of the cross and was assumed to reflect the internal representation of the animation's kinematics.

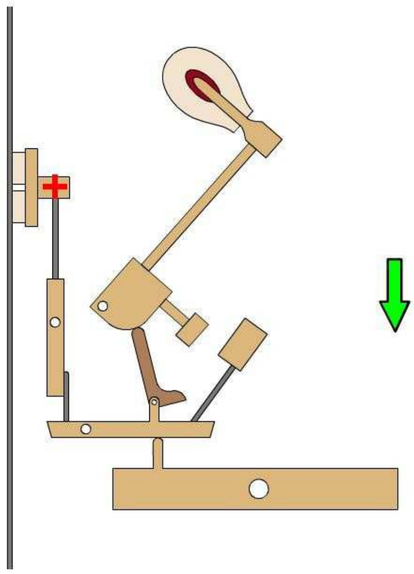
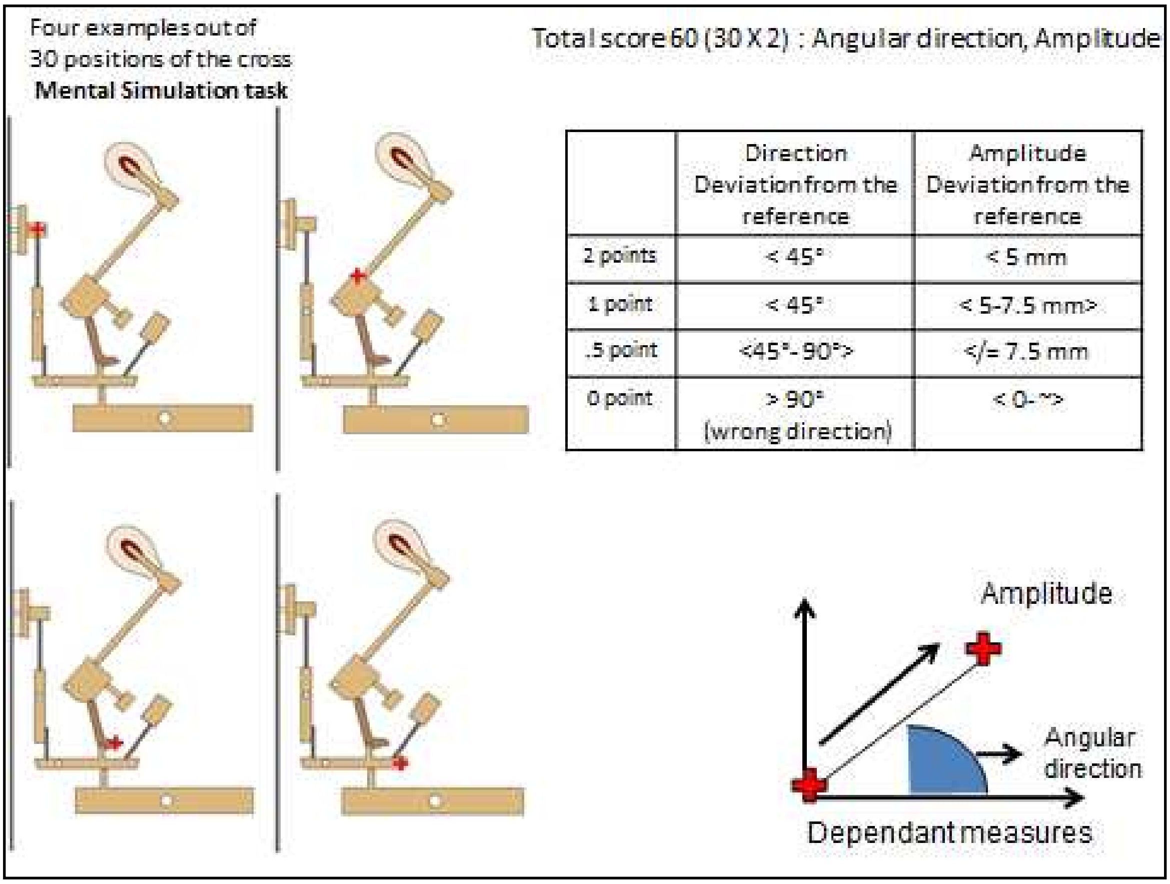


Figure.7.6. Example of the local movement test (position of the cross, taking into the direction of the arrow)



	Direction Deviation from the reference	Amplitude Deviation from the reference
2 points	< 45°	< 5 mm
1 point	< 45°	< 5-7.5 mm>
.5 point	<45°- 90°>	</= 7.5 mm
0 point	> 90° (wrong direction)	< 0-~>

Figure 7.7. Post test-data and indicators of success

After the first post test was completed, participants responded the second post-test. The second post test was the functional mental model test, a comprehension measure of the quality of mental model of the piano operation. Participants had to write precisely the following question:

"What happens with all parts of the piano when a pianist presses the key down and then releases it" Previously we have given a sheet that shows the elements of the piano with their names (Figure 7.8: the elements of the piano). These elements were randomly distributed on the sheet of paper.

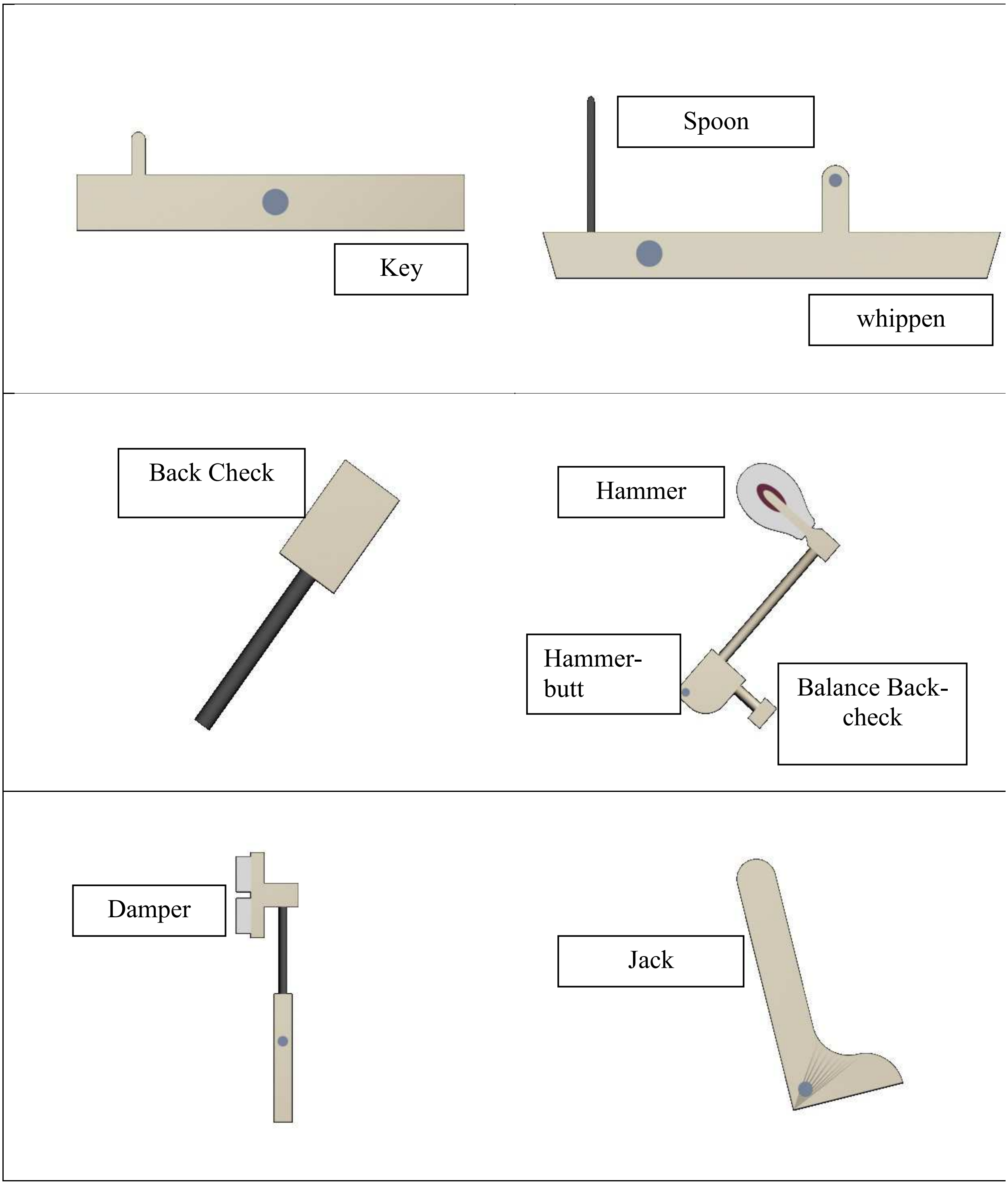


Figure 7.8. The elements of the piano

Scoring guide for functional mental model of the piano operation was based on 15 micro-steps constituting the three main stages of a piano mechanism's functioning (Boucheix and Lowe, 2010) as described on table 7.2. It was used as scoring guide of the second post-test.

Table 7.2 Scoring guide for functional mental model.

SCORING GUIDE POST TEST 2	
	SCORE
Stage 1: Striking	
When the key is pressed, it moves the key-sticker at the end of the key upwards.	1
The key-stricker raises the whippen that makes a rocking motion as a result.	1
The raising of the whippen operates the jack	1
The upward moving jack pushes up the hammer-butt	1
The hammer-butt pivots on its axle	1
The pivoting of the hammer-butt moves the hammer toward the string	1
The hammer strikes the string to produce the note	1
At the same time, the rocking motion of the whippen pushes the damper to lift it off the string	1
The release of the damper liberates the string to sound freely when struck by the hammer	1
Stage 2: Recovering	
The hammer instantly rebounds backwards once it has struck the string	
The balance hammer is caught and blocked by the balance –back check in order to limit the hammer's backward travel. The system stays in this position as long as the key remains depressed	1
Stage 3: Resetting	1
When the key is released, the whippen drops	1
The back-check releases the balance hammer	1
The jack moves downward under the butt and the hammer returns to its initial position	1
At the same time, the damper returns to the string	1
SCORE TOTAL	

The score was assigned as 1 point for correct answers, 0.5 points for incomplete answers and 0 points for wrong answers. This test has been widely used in literature and was considered the most sensitive because it can see through the drafting mental models that have been created.

Results

Figure 7.9 and table 7.3 show the comprehension performance (% of correct answers) for each post-test (local movement and functional mental model).

A repeated measure ANOVA with four presentation modalities as the between subjects factor and the two post test (local movement and functional mental model) as within subjects factor was performed. This analysis showed that the effect of the presentation factor was not significant, F (3, 77) = 1.12, *p* =. 34 but the effect of the post test was significant F (1, 77) = 7, 99, p < .01.

Table 7.3. Mean score and standard deviation (SD) the comprehension performance (% of correct answers) for each post-test (local movement and functional mental model).

Type of conditions				
	St+An (n=22)	An+St (n=20)	SS (n=21)	AA (n=19)
Post test 1 (The Local movement test)	41.78(14.87)	38.83 (13.93)	42.18 (8.01)	45.57 (13.45)
Post test 2 (The Functional mental model)	24.85 (12.86)	20.83 (15.70)	16.03 (9.41)	13.33 (9.73)

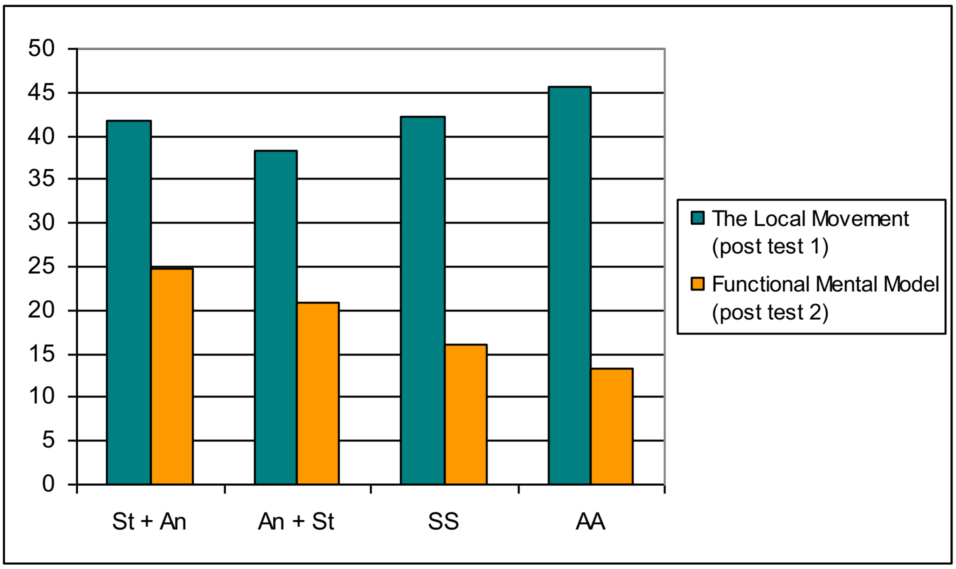


Figure 7.9. Mean score in percent for the local movement test and the functional mental model test in each type of presentation.

The ANOVA showed also a significant interaction between the type of presentation and the post-test, F (3,78) = 6,52 p < .001, η_p^2 = .20. The type of presentation effect was different across the two post-test: in favor of ***St+An*** condition for the mental model measure and in favor of ***An+An*** for the local motion recall task.

Univariate analysis showed a significant effect of the type of presentation for the mental model score, F(3,74) = 4.44, p = .006, η_p^2 = .15. Scores for ***St+An*** and ***An+St*** were higher than scores for ***St+St*** and ***An+An***, F(1,74) = 10,01, p = .002. Univariate analysis for the local motion score showed no effect of the type of presentation, F(3,74) = 0.85, p = .046, η_p^2 = .03.

Finally, local motion scores were higher than mental model quality scores $F(1,74) = 290.05$, $p < .001$, $\eta_p^2 = .79$.

Discussion

In this experiment, we compared two dual format presentations to two single format presentations. Experimental condition 1 (St+An) showed the highest score in both post-test, the local movement test and the functional mental model test. Several explanations are possible for this result. The format of six key steps static pictures seems concise and very clear and allows the subject to identify the key stages of the process. Laird, Tardieu, and Cavazza (1993) presented that simplified mental models could provide the students with a simplified representation of complex dynamic process. It can promote subject understanding how the mechanism works. However, if these simplified models are useful to understand and assign relative positions of different parts of piano (their link with each other) across the different stages of the movement of the upright piano mechanism, they are not sufficient to get an accurate picture of the functional relations between group of components involved in the functioning of the system and sub-systems. The animation following the series of static pictures gave learners the opportunity to test their (multiple) mental models of the functioning of the piano acquired during the first stage with the segmented presentation. With the animation, they could check the series of inferences created when learning from the series of static pictures.

Learners in the experimental condition 2 (An+St) performed at a medium level for functional mental model test but at the lowest level for the local movement test. May be, when the presentation of the animation is in the first position, the learner collects mainly information about the local movement of the piano pieces but not enough on their relations. So may be the learner's attention, in the second stage with the series of static pictures stay more focused on the local movement retrieving than on the complex relation between the events (a component plus its associated behaviour) of the piano mechanism.

The comprehension scores in the experimental condition 3 (St+St) were in third position for the functional mental model test. Participant could create inferences but maybe he could not check them from the feed-back given by an animation. This is showed in poor result of functional mental model test. For to the local movement test, these results are almost identical to those in condition 1. This was predictable because the learner did not see the animation. He could not assess the local movement. This made the comprehension test more difficult for participants.

The comprehension result of the experimental condition 4 (An+An) showed the lowest level for the post test 2. Therefore, we can assume that participants just tended to look at the animation. This tendency resulted in difficulties in making functional inferences about the relation between components and group of components.

Conclusion

As a summary we can conclude that dual format presentation should be better for developing a dynamic mental model than single format presentation. In the next experiment (experiment 2), we will focus on a new format of cueing (temporal cueing) in segmented static pictures which depict a dynamic process. In order to understand more the result of the previous experiment, especially the behaviour of participants, an eye tracking analysis will be used.

CHAPTER 8

EXPERIMENT 2. The use of visual 'tokens' in cueing temporal features of series of static pictures and animation in dynamic process comprehension.

Introduction

Some researchers proposed to improve learning from animations by guiding attention to its most essential aspects with cueing or signaling approach (Mautone & Mayer, 2001). Cueing can be defined as the addition of non-content information that captures attention to those aspects that are important in an animation (e.g. coloring, arrows) (Mautone & Mayer, 2001). Cues are non-content devices such as arrows and colours that are added to visual display to reduce the search space a learner must explore and increase the task-relevant aspects come into the focus of attention. They do not provide new information or change the content of the instructional materials (Lorch, 1989). Mayer (2001) stated that cues are intended to help learners in selecting relevant information, organizing, and integrating the information into a coherent representation. By adding visual cues to a complex animation the visual search associated with locating relevant aspects should be reduced, thereby reducing ineffective cognitive load and allowing more cognitive resources to be allocated to learning (Paas, Renkl, & Sweller, 2003).

Prior research with static representation demonstrated that the effect of cueing have clear implications for conceptual understanding, it may improve the recall of texts, its organization in memory, and may lead to a better understanding of text content (Loman & Mayer, 1983; Lorch, 1989; Lorch & Lorch, 1996). Several studies have shown that cues may improve learning from text and picture (Chandler & Sweller, 1997; Kalyuga, Chandler, & Sweller, 1999) or picture alone (Grant & Spivey, 2003). Using color coding showed the link between textual and pictorial information more clear, improved learners' understanding of the presented information (Kalyuga et al., 1999). Moreover, James, Gavota, & Quarieau (2008) proved that cues can reduce the amount of experienced cognitive load and Tabbers, Martens, & Merriënboer (2004) suggested that cues can effectively increase working memory resources essential for learning. Therefore, guiding learner's attention to relevant aspects in static representations by cueing them increase the possibility that learners extract the right information and allows them to process the information more deeply and more understanding.

However, cueing techniques found to be effective with static pictures are not effective for cueing animation (De Koning & al., 2009). Static cueing methods seem not transferable to animations. The APM (Animation Processing Model) developed by Lowe & Boucheix (2008, 2011) would have predicted this lack of effectiveness with animation. Signaling methods with static pictures were exclusively focused on the cueing of the components, e.g. the entities, of the content depicted and not on the temporal events of the dynamic process. An event can be defined as a component and its associated behaviors, Lowe and Boucheix, 2008. Events are dynamic and have a greater visual saliency than components; even if these components could have a big size. Recently, Boucheix & Lowe, 2010, Boucheix and Lowe, in press, showed that temporal cueing of events (with dynamic signals within the animation) were more efficient than traditional entity cueing.

In the present study, a new form of spatio-temporal cueing of series of static pictures was tested. This new form of cueing consisted in the temporal and sequential cueing of groups of components which are involved temporally across the different stages of the functioning of the dynamic process, or of the mechanism, to be learnt (see figure 8.1). Each group of two, or more, related functional components could be called a “token”. In the present study, tokens were cued temporally. As in the previous study, dual format presentation, a series of static pictures and an animation was used. Tokens were cued temporally in the two formats. It was expected that this technique of spatio-temporal cueing of tokens would result in better comprehension performance than a no-cued presentation.

In order to study how the cued tokens were cognitively processed we used an eye-tracking method during learning.

Research Goals

The main purpose of this experiment was to study the cognitive process in learning a complex mechanism system with cueing spatio-temporal tokens. This experiment investigated: (1) dual format presentations (a series of static pictures and an animation format presentation with cueing would result in better comprehension than single format presentations with cueing also. (2) Whether using spatio-temporal tokens would result in better comprehension scores compared to a control no-cued presentation, with dual format modality.

(3) To complement the main goal, the effect of executive functions on the understanding of animation was also tested. Four short tests were used in this experiment to evaluate the possible effect of working memory capacities on the understanding of the complex mechanism system from animations.

Hypothesis

In this experiment, two formats of presentation (static sequential and animation) were viewed in four conditions: (1) Sequential-Animated/**SA**, (2) Animated-Sequential/**AS**, (3) Sequential-Sequential/**SS** and (4) Animated-Animated/**AA**. In relation to the result of the previous experiment (experiment 1), dual format should be better for developing a dynamic mental model functional than single formats presentations.

An effect of attention guidance for building of mental model should give higher point for comprehension score in all conditions than the result obtained in Experiment 1 (Hypothesis 1). In order to test precisely this hypothesis in experiment 2, the dual format cued conditions were compared to a dual format no-cued condition.

Animations processing could need a high level in speed processing, a high level in the activity of refreshment of current representations, and a high level in the inhibition of passed pictures during learning. We expected a positive and significant correlation between the level of executive functions and the performance in comprehension (Hypothesis 2).

Regarding eye tracking measures, it was also predicted that in the different conditions, learners should exhibit different patterns of attention according to the presentations formats. More precisely, temporal tokens should efficiently direct learner's attention on the group of components involved across time along the different stages of the piano functioning (e.g. respectively the striking stage, the rebounding stage and the resetting stage), (Hypothesis 3).

Method

Participants

Eighty nine students from Bourgogne University took a part as participants in the experiment. The design of experiment is shown in table 8.1. They were all over 18 years old and their average age was 20 years (SD=1,9) with 81% females (which is a common percent in the French psychology departments of Universities). They were all undergraduates' students in psychology and they are given 0.5 credit point for their participation. Previously, participants has verified that they did not know how a piano mechanism working.

Table 8.1. Design of experiment 2, with four groups

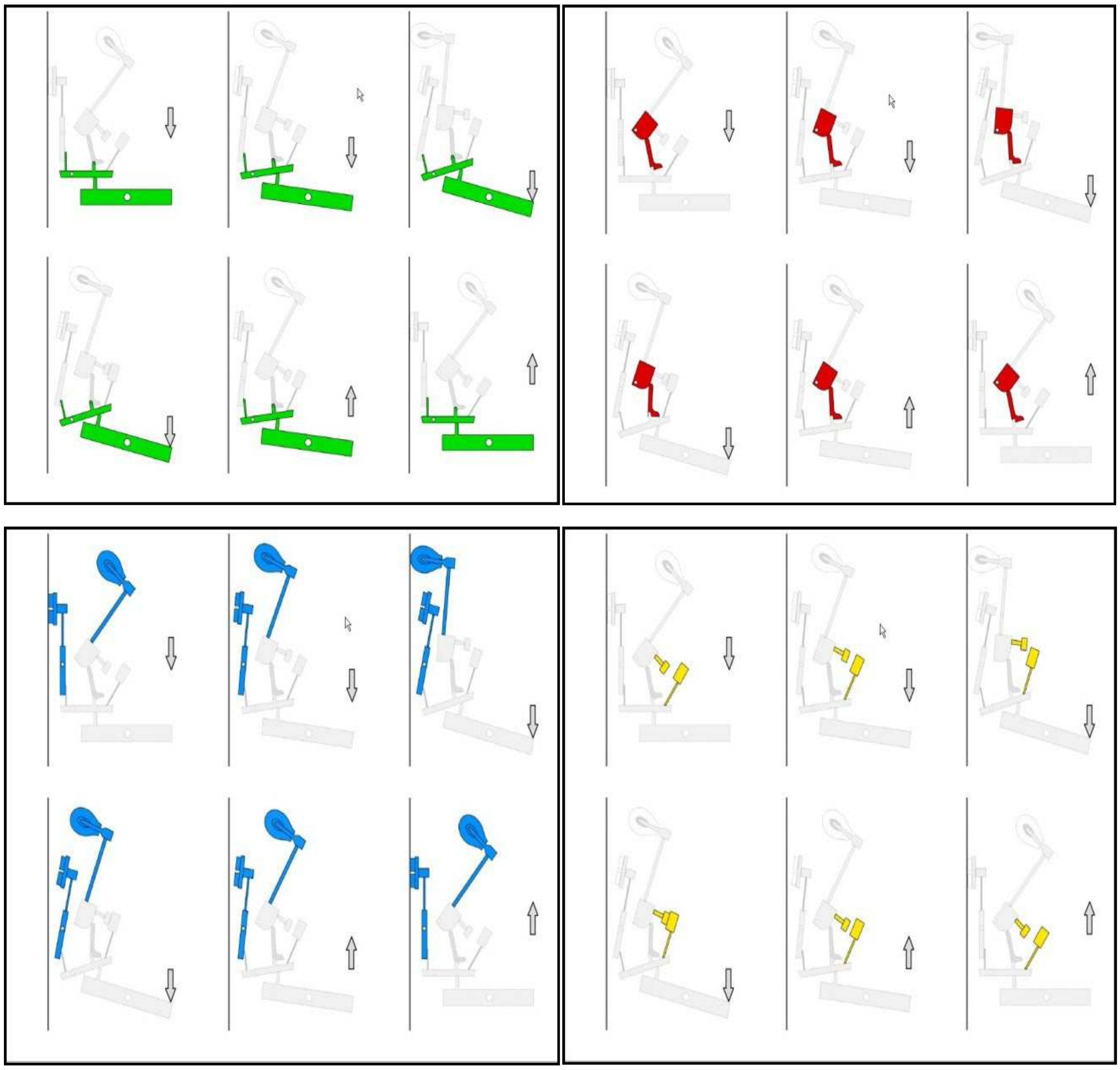
SECOND PRESENTATION	FIRST PRESENTATION	
	Static (S)	Animation (A)
Static (S)	SS (N=18)	SA (N=18)
Animation (A)	AS (N=18)	AA (N=18)
Control	Control, only AS (N= 17)	

Material

Series of six static pictures presentation

Similarly with previous experiment, we used a series of six static pictures which represented six pictures of the key steps of the operation of an upright piano mechanism. We designed four tokens presented sequentially with different colours in the series of static pictures (figure 8.1): token 1 cued the key and the whippen in green; token 2 cued the jack and the hammer-butt in red; token 3 cued the hammer and the damper in blue: and token 4 cued the balance and back-check in yellow). The series of six static pictures with tokens in colour was presented in replicated view in four times with presentation time of 26 seconds per screen. Each presentation has a single token colour and the rest of the display was a neutral slight brown colour, that is to say anti-cued. The order of presentation was consistent with the progressive changing dynamics of the operation piano mechanism. The first screen was token 1 key-whippen (green), then hammer butt-jack (red), followed by damper-hammer (blue), and balance-back check (yellow) was the last display (Figure 8.1). Overall duration of the sequential presentation was one minute and 45 seconds, as in experiment one.

1 2



3 4

Figure 8.1. Type of sequential static pictures presentation with four tokens

Note: Token 1 (Key-Whippen) in green, Token 2 (Hammer butt-Jack) in red, Token 3 (Damper-hammer) in blue, Token 4 (Balance-Back check) in yellow.

Animation presentation

The animated presentation of the upright piano mechanism has been created with the same cued tokens as in the sequential static pictures presentation, (Figure 8.2).

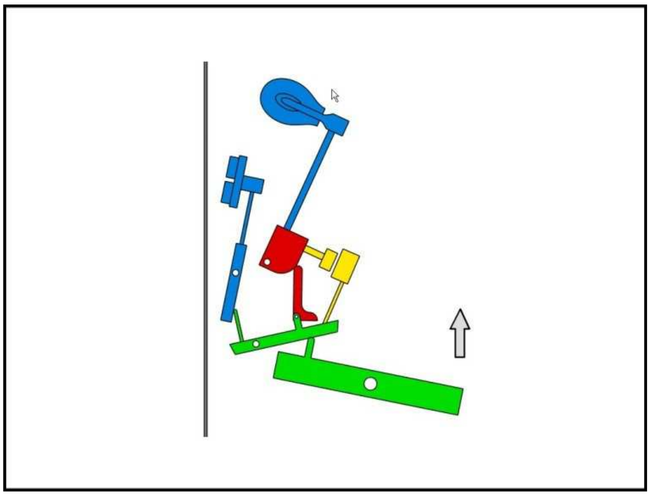


Figure 8.2. Cued animation presentation

For the presentation of the animation, tokens were cued progressively and chronologically along the three phases (Strike phase, Rebound phase, and Reset phase) which was based on how does a piano mechanism works (Boucheix and Lowe, 2010). These phases have been used for the scoring guide of the functional mental model test. Animation was presented in five times for 20 seconds. Total duration of the animation presentation was one minute and 45 seconds. There was a sound that happened when the hammer hit the string and visual signal or highlighting previewed when the hammer hit the string and when back catch contact to back check or the hammer rebounded the string. Participants have no control over the speed of the animation. As in experiment one, participants might have access to look the names of the elements of piano with removable windows or pop up systems.

Control condition

The presentation of the material in control condition was similar to previous experiment (Experiment 1), with a representation of the piano mechanism in a full brown colour (Figure 8.3). Only one of the four possible conditions of presentation of the dual format visualization was used as control condition: the St+An, in which the series of static pictures was presented first and followed by the animation. The dual format presentation was applied in the same duration time as the other condition of the experiment.

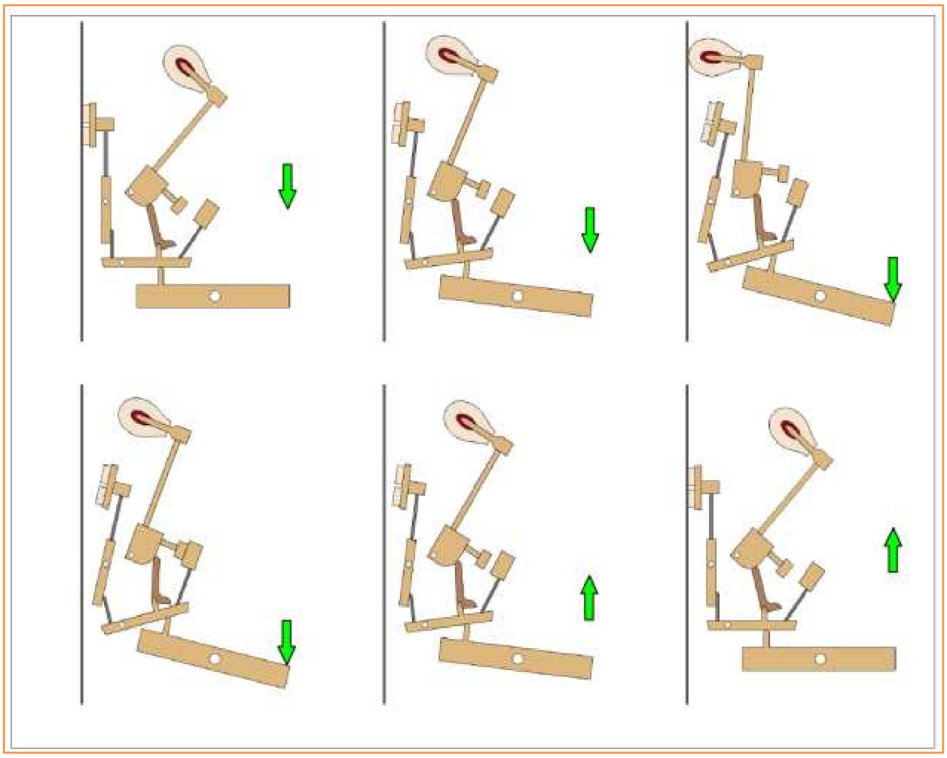
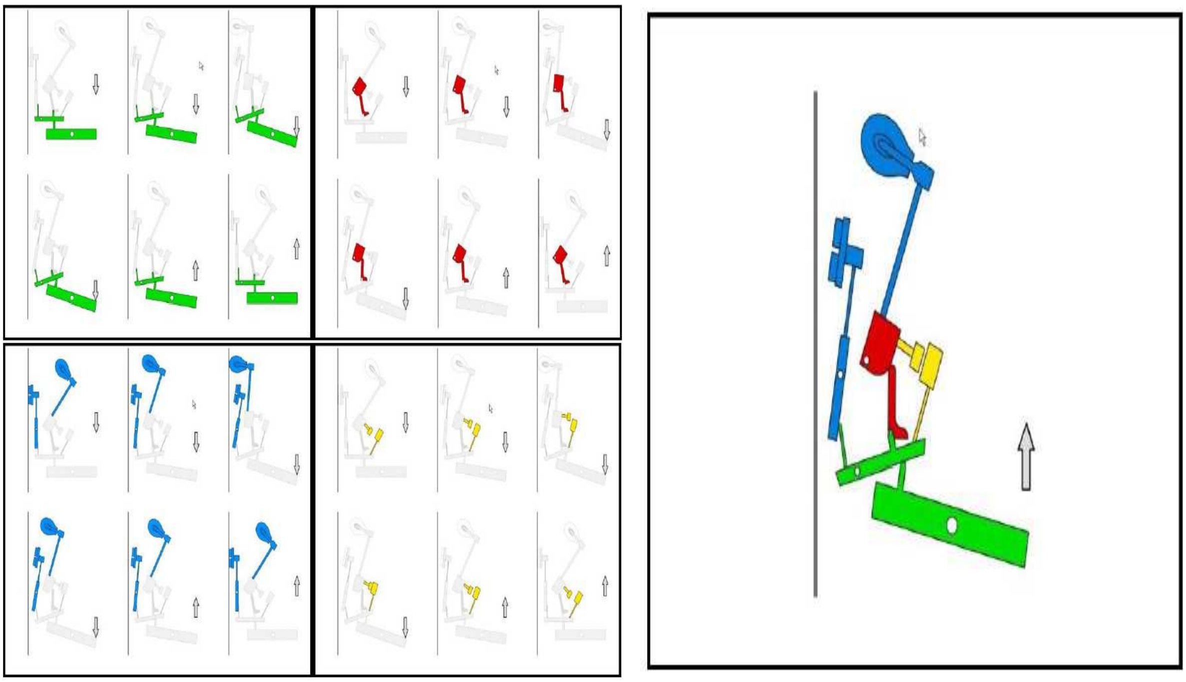


Figure 8.3. Type of control condition

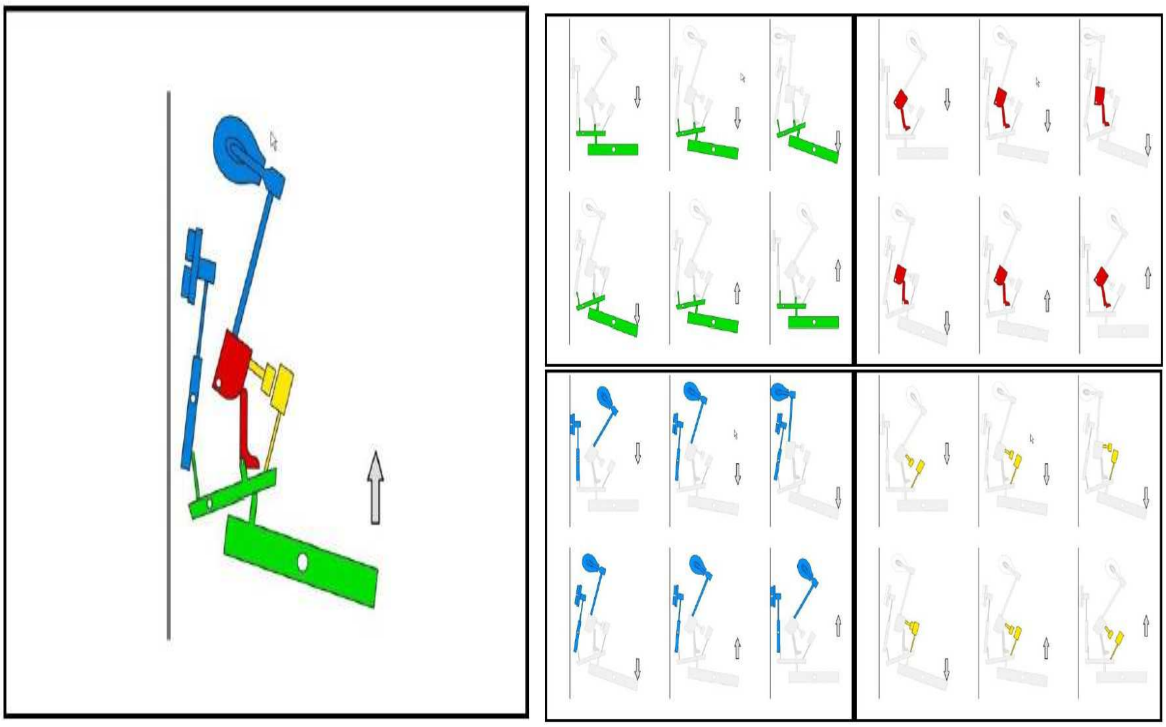
Experimental condition

The experiment was designed in five experimental conditions: (1) the series of static pictures, a sequence with spatial cueing (tokens colored), followed by the animation (with the sequence of colored tokens) (SA- Figure 8.4), (2) the animation (sequential presentation with tokens) followed by the series of and static pictures, the sequence with spatial cueing (AS-Figure 8.5), (3) the series of static pictures, in sequence with cueing, followed by the same series of static pictures with spatial cueing (SS- Figure 8.6), (4) the animation (with the sequence of colored tokens) followed by the same animation (with the sequence of colored tokens) (AA- Figure 8.7), and (5) the series of static pictures without cueing (presented in sequence but remaining on the screen simultaneously) followed by the animation presentation without cueing (Control condition- Figure 8.8).



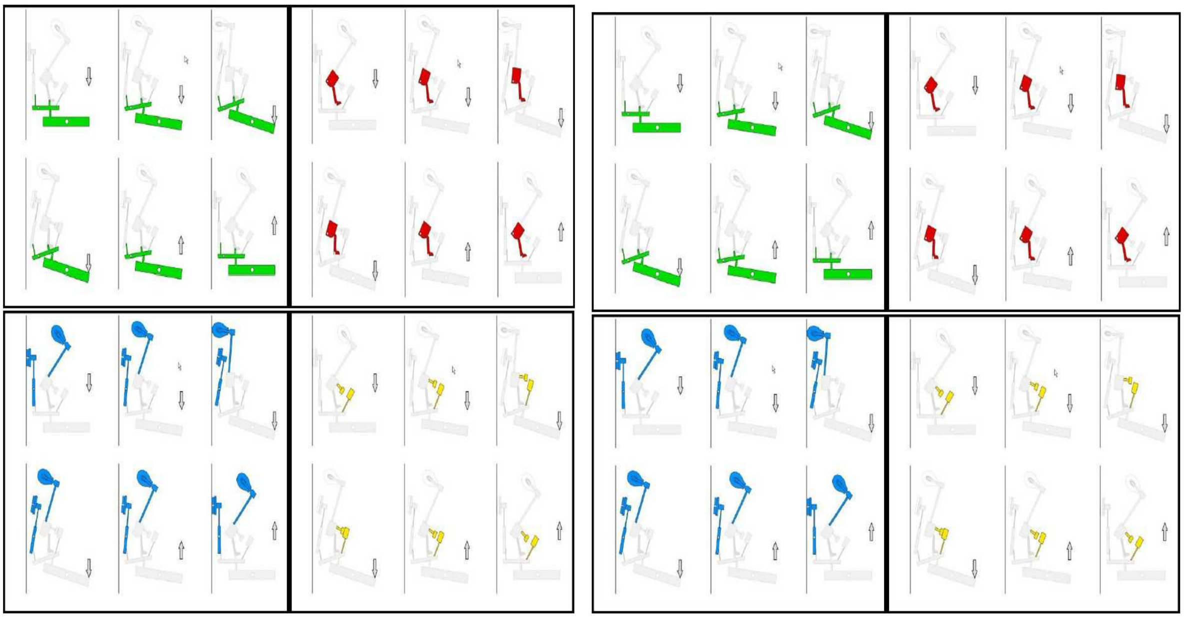
1 minute 45 second　　1 minute 45 second

Figure 8.4. SA condition



1 minute 45 second　　1 minute 45 second

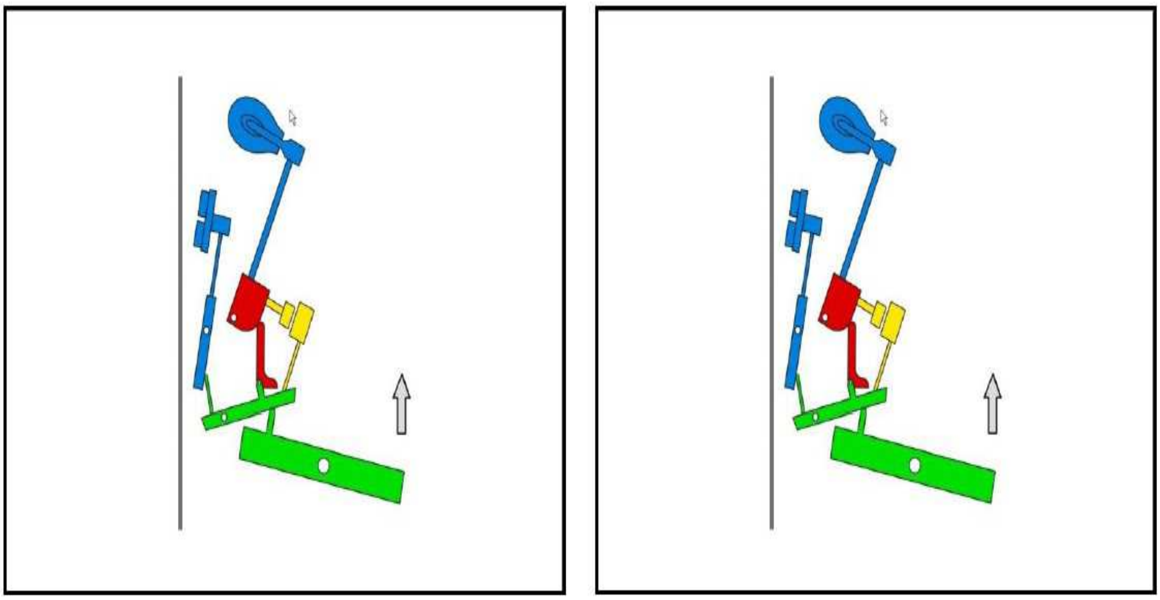
Figure 8.5. AS condition



1 minute 45 second

1 minute 45 second

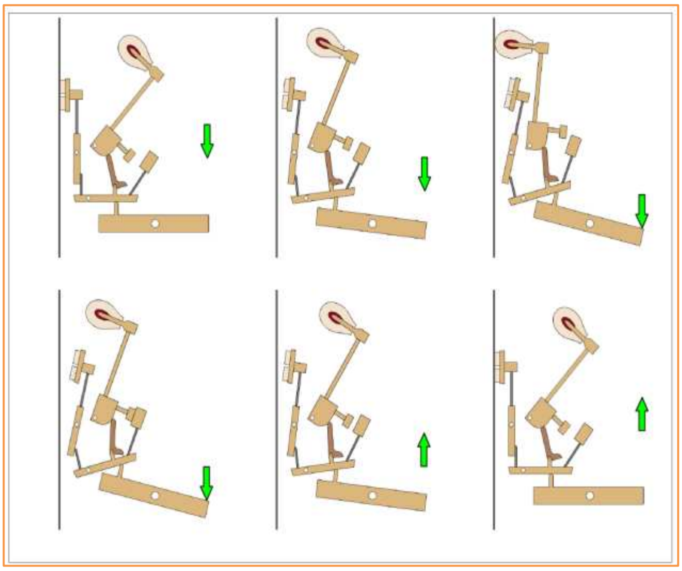
Figure 8.6. SS condition



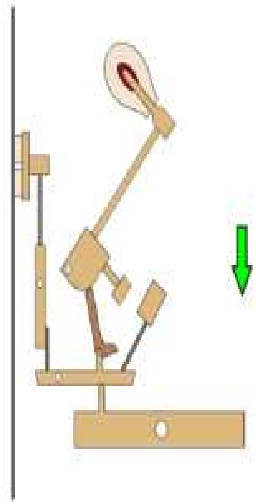
1 minute 45 second

1 minute 45 second

Figure 8.7. AA condition



1 minute 45 second



1 minute 45 second

Figure 8.8. Control condition

Eye movement recording and data analysis

Eye movements were recorded with a 120 Hz Tobii, 1 binocular corneal reflectance and pupil centre eye tracker (Tobii Technology, Sweden, Stockholm). Data were recorded with Tobii-Studio software. The computer screen (17'') for displaying the animation was positioned 60 centimeters from the participant. The binocular tracker had a field of view of approximately 20 cm x 15 cm x 20 cm (width x height x depth) and an effective tolerance for free head-motion of about 30 cm x 15 cm x 20 cm with an accuracy of 0.5°.

Four areas of interest (AOIs) were created. AOI's were defined for both formats of the dual presentation, the series of static pictures and the animation. These four AOIs were in correspondence with the four tokens: AOI-1 is token 1 (Key-Whippen) in green, AOI-2 is token 2 (Hammer butt-Jack) in red, AOI-3 is token 3 (Damper-Hammer) in blue, and AOI-4 is token 4 (Balance-Back check) in yellow. These AOIs corresponded to a functional part of the piano system (Figure 8.9).

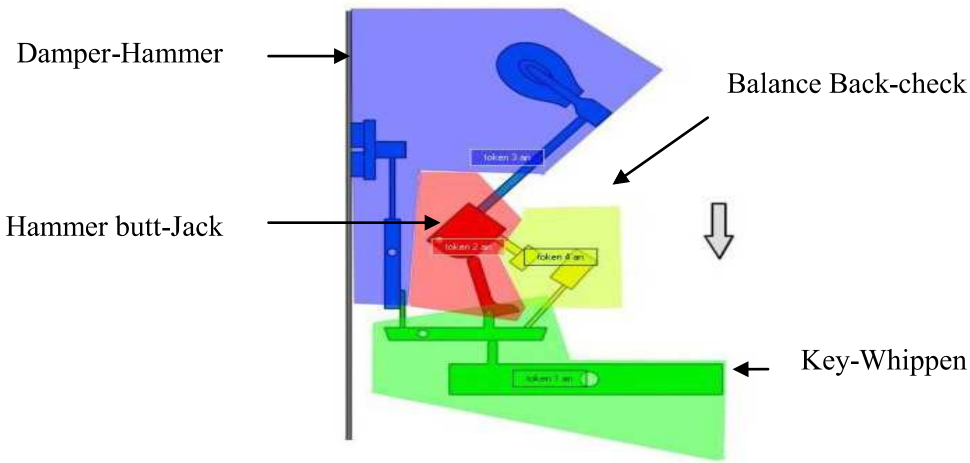
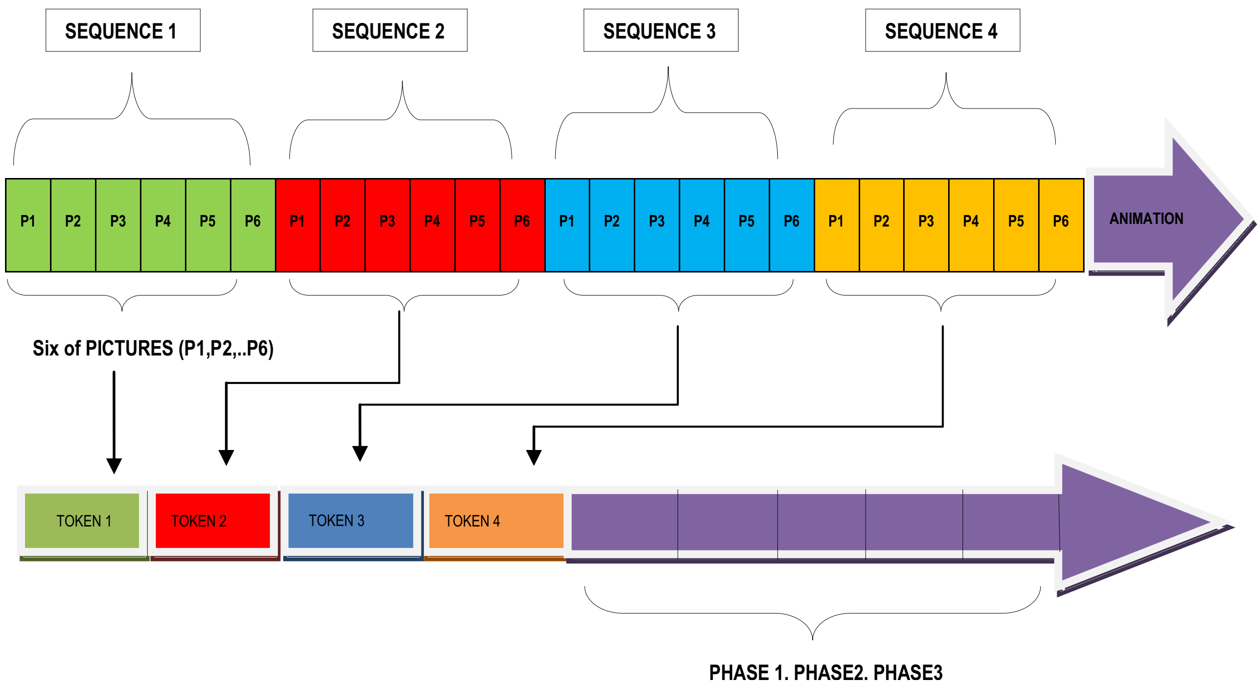


Figure 8.9. Four area of interest (AOIs) in an animation presentation

Three types of common measures were used for the analysis of the eye movement data for each presentation format: the fixation duration (the total fixation duration in seconds within an AOI), the observation count (number of visits to an AOI), and the number of fixation before the arrival in an AOI (the number of fixation before the participants fixated within an AOI for the first time). The fixation filtering threshold was set at 100 ms.

Eye movements (for each AOI) of each participant were analyzed in time segments across each learning step presentation, as shown in Figure 8.10. Such analysis is a time locked investigation.

Series of STATIC pictures 1’45 “



ANIMATION 1’45’’

Figure 8.10. Time segments in for each step of the learning time used for the eye movement analysis (example: SA condition)

Figure 8.10 shows that the eye movement analysis was composed of four sequences of operation of the upright piano mechanism. Each sequence contained a series of six static pictures. For each sequence we had six static pictures with a single colored token (Figure 8.1). The order of presentation was consistent with the changing dynamics of operation piano mechanism. There were three phase in the animation (Strike phase, Rebound phase, and Reset phase).This segmentation of the data was used for the eye movement analysis.

Comprehension test and data analysis

Comprehension was measured using by two-post tests which were similar to those in the previous experiment (Experiment 1). The first post test was the local movement test (see page 90, Figure 7.7.) and the second post test was the functional mental model. Scoring criteria for functional mental model test is described in table 8.2. One point was given to a participant for each correct written description of an event, for each token in each phase of the piano mechanism. Two types of scores were calculated, one by phase, one by token as shown in table 8.2.

Table 8.2. Score guiding for functional mental model test.

SCORING GUIDE POST TEST 2	
PHASE 1- Striking	SCORE
Token 1 When the key is pressed, it moves the key-sticker raises and to the extreme left The movement of the key makes the whippen up that makes a motion left and right up	 1 1
Token 2 The raising of the whippen pushes up the jack The jack pushes up the hammer-butt The hammer-butt pivots on the left	1 1 1
Token 3 The movement of the hammer-butt moves the hammer toward the string (left) The hammer strikes the string and produces the sound At the same time, the rocking motion of the whippen and the damper to left off the string The release of the damper allows the string to vibrate and produce the sound	 1 1 1 1
PHASE 2-Rebounding	
Token 4 After the hammer strikes the string, he rebounds backward The balance hammer is blocked by the balance back-check in order to limit the hammer's backward The system stays as long as the key remains depressed	 1 1 1
PHASE 3-Resetting	
Token 1 When the pianist releases the key, it returns to the horizontal, the same as with whippen **Token 4** The back-check releases the balance hammer **Token 2** The jack moves downward under the hammer-butt and the hammer returns to its original position **Token 3** At the same time the damper back against the string and the hammer returns to its original position on the right	 1 1 1 1

Token 1	Token 2	Token 3	Token 4	***TOTAL***
/3	/4	/5	/3	***/15***

Spatial ability test

Spatial Relation test (Booklet 4) form of the differential aptitude test (DAT5, Bennett, Seashore and Wesman, 1973, 2002) was used to measure spatial ability for each of participant. This test consists of thirty items tasks the capabilities of processing and mental rotation from geometric shapes. The participant has fifteen minutes to choose the right solution for complete thirty multiple choice questions. One an example of items spatial relation test is shown in Figure 8.11. One point is given for each correct answer and maximum total correct scored is 30 points. This test was used for the control and checking the equality of spatial abilities performance in each experimental group.

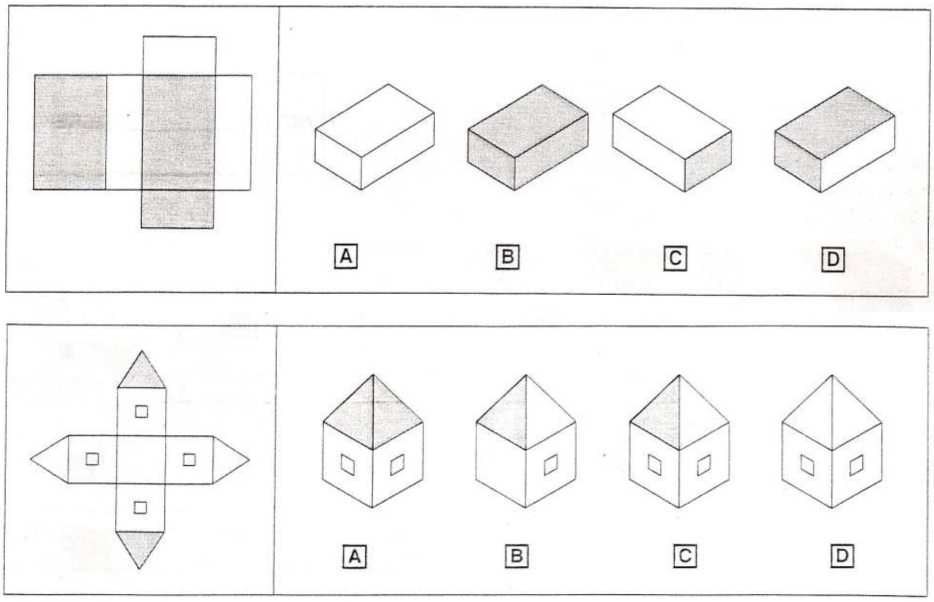


Figure 8.11. An example of items spatial relation test (Booklet 4, DAT5, Bennett, Seashore and Wesman, 1973, 2002)

Executive function test

We used four executive function test: (1) The Stroop test, for testing the capacity of inhibition of each participant (2) the N-Back test for the evaluation of the abilities to maintain and update information in working memory, (3) the XO test, for the measure of speed processing and (4) the Plus-Minus test, for the evaluation of attentional switching from one task to another and to investigate cognitive flexibility. For more detail description about executive function test, we reviewed those in these following:

Stroop test

John Ridley Stroop introduced the Stroop effect in 1935. The Stroop effect is a measure of the reaction time of a task. In the Stroop test, there were three different kinds of stimuli. In the first one, the names of colours appeared in black ink. In the second, there were squares of given different colours. In the third, the names of colours appeared in a different ink than the colour named (Figure 8.12).

Stimulus 1: **Purple Brown Red Blue Green**

Stimulus 2: 

Stimulus 3: **Purple Brown Red Blue Green**

Figure 8.12. Examples of the three stimuli and colors used for each of the activities of the original Stroop article (Stroop, 1935)

In the Stroop task, the subjects are asked to name the colour that colour words are printed in when the ink colour and word meaning often conflict (for example, the word "Purple" in red ink). Stroop identified that subjects need more time to complete the colour reading in the third stimulus compared to the naming of the colour of the squares in the second stimulus. Executive Function are needed to performed this task, as the relatively over learned and automatic behaviour (reading a word) has to be inhibited in favour of a less practiced task-naming the ink colour.

N-Back test

N-back task is a continuous performance task that is used in neuropsychology. This test was introduced by Kirchner, 1958. The participant is presented with a sequence of stimuli and the task consists in indicating when the current stimulus matches the one from *n* steps earlier in the sequence. In the example presented in the diagram below, the experimenter will read out several letters sequence, the subject try to listen carefully and memorize it, then give the respond to say 'yes' or 'no' if he/she thinks that each new letter is the same as with two letters before. In other words, when the experimenter read out "Z", "R" and "K", is that "K" is the same with two letters before ("Z")? The subject should respond "NO". However, if then the experimenter read "R" after "K", the subject should respond "YES" because "R" is the same with the second letters before.

Example:	Z – R – K	Z – R – K – R
	⇖ ⇗	⇖ ⇗
	NO	**YES**

Speed test (X-O comparison test)

X-O comparison test is the letter comparison test (Salthouse, 1990). Participant is presented a page containing pairs of letters (X-0). The subjects are instructed to decide whether the two members of the pair were identical or different and give the tick according to the 'identical' or 'different' column. They were given 30 seconds to make as many comparisons as possible. The score was the total number of the correct comparisons.

COMPARISON XO

		IDENTICAL	DIFFERENT
X	O	☐	☐
X	X	☐	☐
O	X	☐	☐

Plus-minus test

In the Plus-Minus test, participants have to complete each list of presented numbers as quickly and accurately as possible without returning back, see the diagram below. The instruciorn are as follow: for list A, add the number 3 to each of the numbers; for list B, subtract the number (-3) at each of the numbers, and finally, for list C, add 3 to the first number, subtract 3 for the second, and add 3 for the third, and continue for all the numbers on below.

List A		List B		ListeC	
+3		**-3**		**+3, -3...**	
18	**.....**	52	**.....**	27	**.....**
67	**.....**	94	**.....**	16	**.....**
31	**.....**	37	**.....**	40	**.....**

Procedure of the experiment

Participant sat in front of the screen at distance approximately 60 cm and the eye tracking system started to calibrate eye movement for each participant before the experimentation started. The overall duration of each experimental condition was 3 minutes and 30 Seconds.

After completing the experiment, participant continued to carry out post test (comprehension test, spatial ability test, and executive function test). The total duration of each participant for doing our experiment was approximately one hour. Participant received credit point as reward to participate in this experiment.

Result

Comprehension test

The following table showed result of Mean in percent and (SD) for the first post test (local movement test) and the second post test (Functional mental model).

Table 8.3. Mean score in percent and (SD) Post test 1 and Post test 2

	Type of Conditions, presentation format				
	SA (n=18)	AS (n=18)	AA (n=18)	SS (n=18)	Control (n=17)
Post test 1 (Local movement test)	46.25 (7.63)	47.73 (9.95)	49.31 (9.98)	45.83 (12.02)	50.34 (13.80)
Post test 2 (Functional mental model)	38.89 (13.67)	40.56 (16.62)	42.96 (15.21)	35.19 (17.98)	35.29 (12.53)

For visuo-spatial abilities, the median scores on the DAT based on median split were calculated on all subjects. The median score was 18 out of 30. Participants were divided into two groups for the execution of statistical analysis. Subject with a score below 18 being considered as having low visual-spatial skills (average score was 12.46) and those with scores greater than or equal 18 being considered as having high visual-spatial skills (average score was 22.60).

A one-way ANOVA, with spatial ability as between factor, showed that there was not overall significant difference between the mean scores obtained by DAT in the different experimental group, $F(4,84) = 1.39$, $p = 0.24$.

For comprehension scores, a repeated measures ANOVA, with the two post-test as dependent variables, five experimental conditions as between factor and group level in spatial abilities as categorical factor, showed no significant effect of type of presentation format, $F(4, 79) =$

1.20, p = 0.31 η_p^2 = .057 and no significant interaction between type format presentation and type of post test, F (4, 79) < 1, ns. There was no significant difference in performance between the group in the SA condition and the group in the control condition for the post test-one and post test-two, F (1,79) < 1, ns.

A significant effect on the type of post test was found, F (1, 79) = 33, 34, p < 0.01, η_p^2 = .29, which indicated that the scores of local movements post test were greater than the scores of the functional mental model post test. The interaction between the type of presentation format and the level of spatial abilities (DAT) was significant, F (4, 79) = 2.68, p < 0.04, η_p^2 = .12.

The results of post test 2 by tokens (functional mental model test by tokens) and phases of the piano mechanism are presented in table 8.4.

A repeated measures ANOVA was conducted with scores of each token as within group factor and the conditions of presentation and group level visual-spatial skills as two between factors.

There was no significant effect of experimental conditions, F (4, 79) = 1.04, p = 0.39, η_p^2 = .050 and a significant effect of level of spatial abilities, F(1,79) = 15.48, p<0.0001, η_p^2 = .16.

The interaction between group spatial abilities level and the experimental conditions was also significant, F (4, 79) = 3.20, p = 0.02, η_p^2 = .14. Post hoc test showed mainly that the performance on all subtests for participants with low skills in the group AS, SS, and AA condition were lower than the performances of participants with high spatial abilities. However, for the participants with high level in spatial abilities, no significant differences could be revealed between the different presentation format and particularly between the SA group SA and the control condition group (p = 0.89 and p = 0.86).

Table 8.4. Means scores in percent and (SD) result of Post test 2

Type of Conditions					
	SA (n=18)	**AS** (n=18)	**AA** (n=18)	**SS** (n=18)	**Control** (n=17)
Token 1 (Key-whippen)	46.3 (21.81)	56.48 (32.41)	54.63 (29.6)	51.85 (24.85)	48.04 (24.21)
Token 2 (H. butt-jack)	32.64 (26.48)	22.92 (17.28)	29.17 (22.69)	18.06 (20.66)	19.11 (20.78)
Token 3 (Damper-hammer)	47.78 (16.29)	49.44 (23.63)	53.89 (18.52)	46.11 (27.47)	48.23 (17.40)
Token 4 (Balance-backcheck)	25 (25.08)	33.33 (33.82)	30.56 (24.42)	23.15 (26.9)	22.55 (19.49)
Phase 1 (Striking)	37.04 (15.6)	41.05 (15.38)	45.68 (22.1)	33.33 (20.7)	36.27 (11.96)
Phase 2 (Rebounding)	27.78 (25.57)	38.89 (39.5)	38.89 (31.18)	18.06 (28.19)	33.82 (29.23)
Phase 3 (Resetting)	48.61 (21.39)	39.58 (23.97)	37.5 (20.11)	47.92 (26.86)	33.82 (24.11)

Pos hoc analysis revealed that token 1 and 3 had significantly higher performance than token 2 and 4, $p < 0.001$, no interaction with the type of format presentation, $F(12,237)$, $p<1$, ns.

Another repeated measure ANOVA was conducted with scores of each phase as within group factor and the conditions of presentation and group level visual-spatial skills as two between factors.

This analysis showed mainly that there was a significant effect of the phase; the score of phase 2 was lower than those obtained for phase 1 and phase 3. Score phase 2 was significantly lowest in SS condition (18.06%) compared to AS (38.89%, $p<0.03$) and AA conditions (38.89%, $p<0.03$) and no significant for SA condition (27.78%, $p = 0.29$) and control condition (33.82%, $p = 0.10$).

Eye movement results

As explained above, page 103, eye movement analyses were realized for each format of the dual presentation, the series of static pictures on one side and the animation on the other side. The same analyse was done for each experimental condition. Three types of measures were used: total fixation duration in each AOI, total observation count in each APOI, and the number of fixation before the arrival in each AOI. This last measure was very useful to investigate eye movement when the token appeared on the screen, and particularly cue obedience. Th same segmentation was performed for the control condition.

Results for the series of static picture in type of presentation AS, SA, SS, AA

Result for fixation before, fixation duration length, and observation count in AOI, and in each experimental condition (AS, SA, SS, AA) included the control condition are showed in Figure 8.13, Figure 8.14, and Figure 8.15.

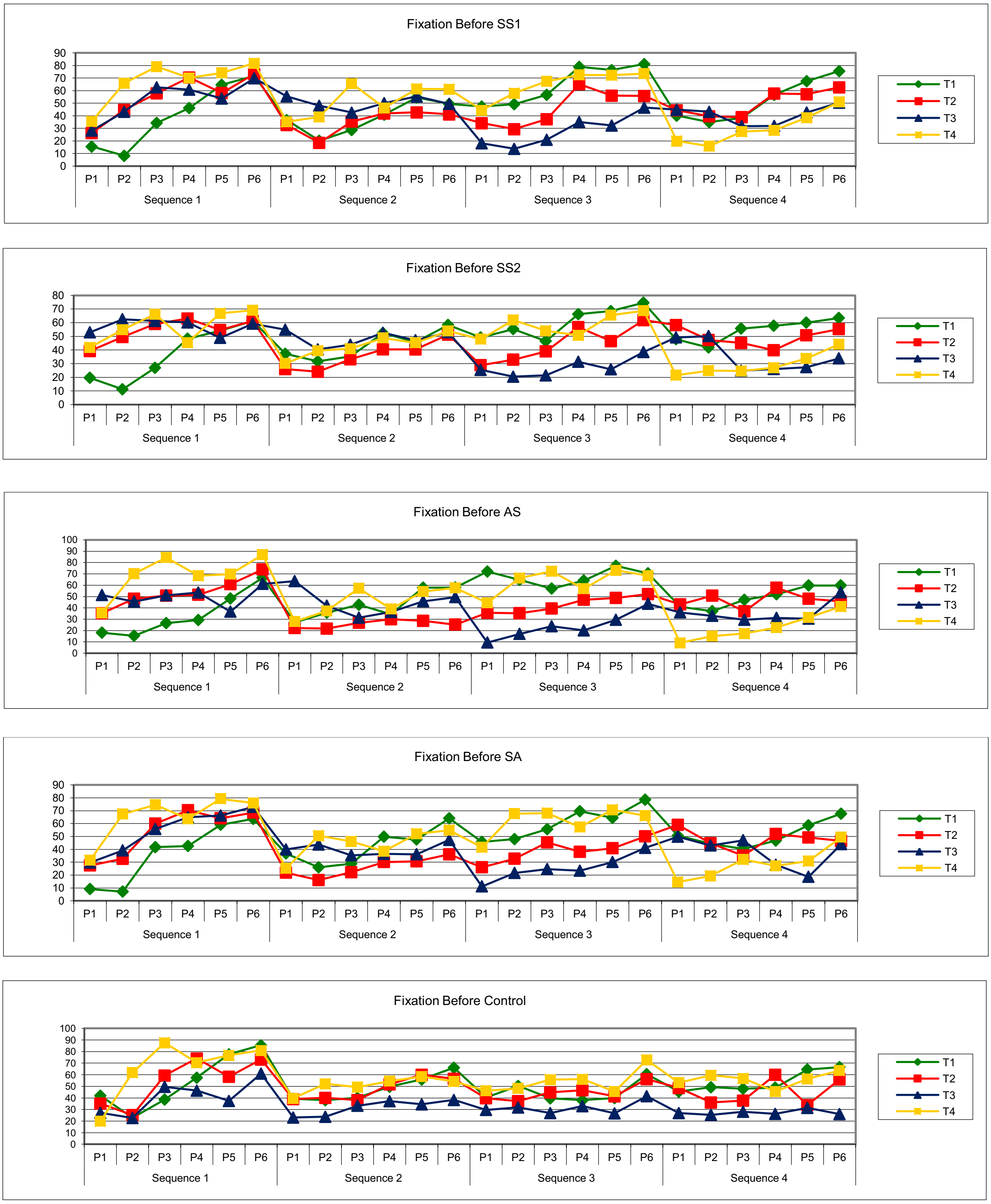


Figure 8.13. Mean number of fixation before the arrival in each AOI (each token); for each experimental condition (AS, SA, SS and AA and control); for each sequence (S), each picture (P1 to P6), and for each token (T1 toT4).

Note: Token 1 (Key-Whippen), Token 2 (Hammer butt-Jack), Token 3 (Damper-hammer), Token 4 (Balance-Back check)

A repeated measured ANOVA on number of fixation before the arrival in each AOI (that is to say each token) was performed with presentation format (AS, SA, SS, AA and Control) as between factor, sequence number, and tokens (AOIs) as within factors. There was no overall significant effect of the different presentation formats (AS, SA, SS, AA and Control), (4, 84) = 0,6 p = 0.66, η_p^2 = .02. The effect of type token was also significant F(3,252) = 34.14, p = 0, η_p^2 = .28, showing that participants arrived sooner in token one. The effect of sequence was significant, F (3,252) = 49.54, p = 0, η_p^2 = .37. Interaction between type of token (AOI) and experimental condition was significant, F (12,252) = 2.39, p = 0.06, η_p^2 = .10. These results indicated that participants arrived sooner in a token (at the beginning of a sequence when the cue appeared) in the AS, SA, SS, AA condition than in the control condition. It means that cue obedience was very effective. Participants fixated the cued tokens when the cue appeared at the beginning of the sequence. The interaction between the sequence and presentation format condition (AS, SA, SS, AA and control) was also significant, F (12,252) = 1.89, p = .03, η_p^2 = .08.

A similar analysis for the duration was performed on results of figure 8.14. A repeated measured ANOVA with the experimental conditions (AS, SA, SS, AA and control) as between factor, sequences and tokens (AOIs) as within factors was performed. There was no overall significant effect of the presentation format (AS, SA, SS, AA and control) because participant spend the same overall amount of time fixating the pictures (animated or static) F(4,84) = 1,16 p = 0.33, η_p^2 = .97. The effect of sequence was significant, F (3,252) = 10.25, p = 0, η_p^2 = .10, less time was spent on sequence two (which is less salient) compared to sequences 1 and 3. The effect of type of token (AOI) was also significant, F (3, 252) = 25.72, p = 0, η_p^2 = .23 showing that participants spent also less time on token two than on the other tokens. Interaction between the effect of type token and experimental condition was significant, F (12,252) = 2.5, p = 0.04, η_p^2 = .10. In particular, compared to the control condition participants in the experimental conditions (AS, SA, SS, AA) spent more time on a cued token (compared to the other token un-cued at this time) when this token appeared at the beginning of a sequence. This result also means that participants followed and used the cues.

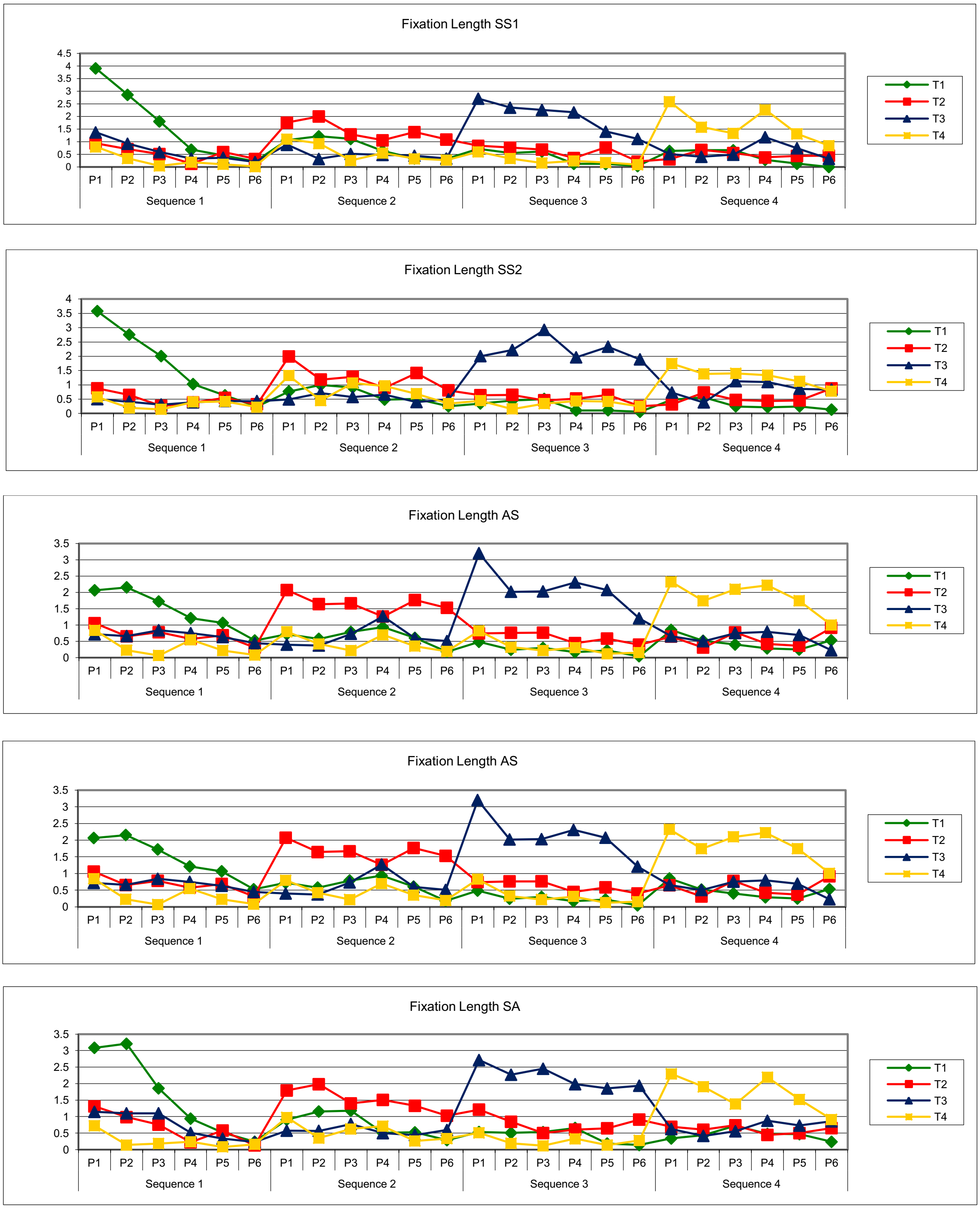


Figure 8.14. Mean fixation length (duration) in each AOI (each token), for each experimental condition (AS, SA, SS and AA and control); for each sequence (S), each picture (P1 to P6), and for each token (T1 toT4)

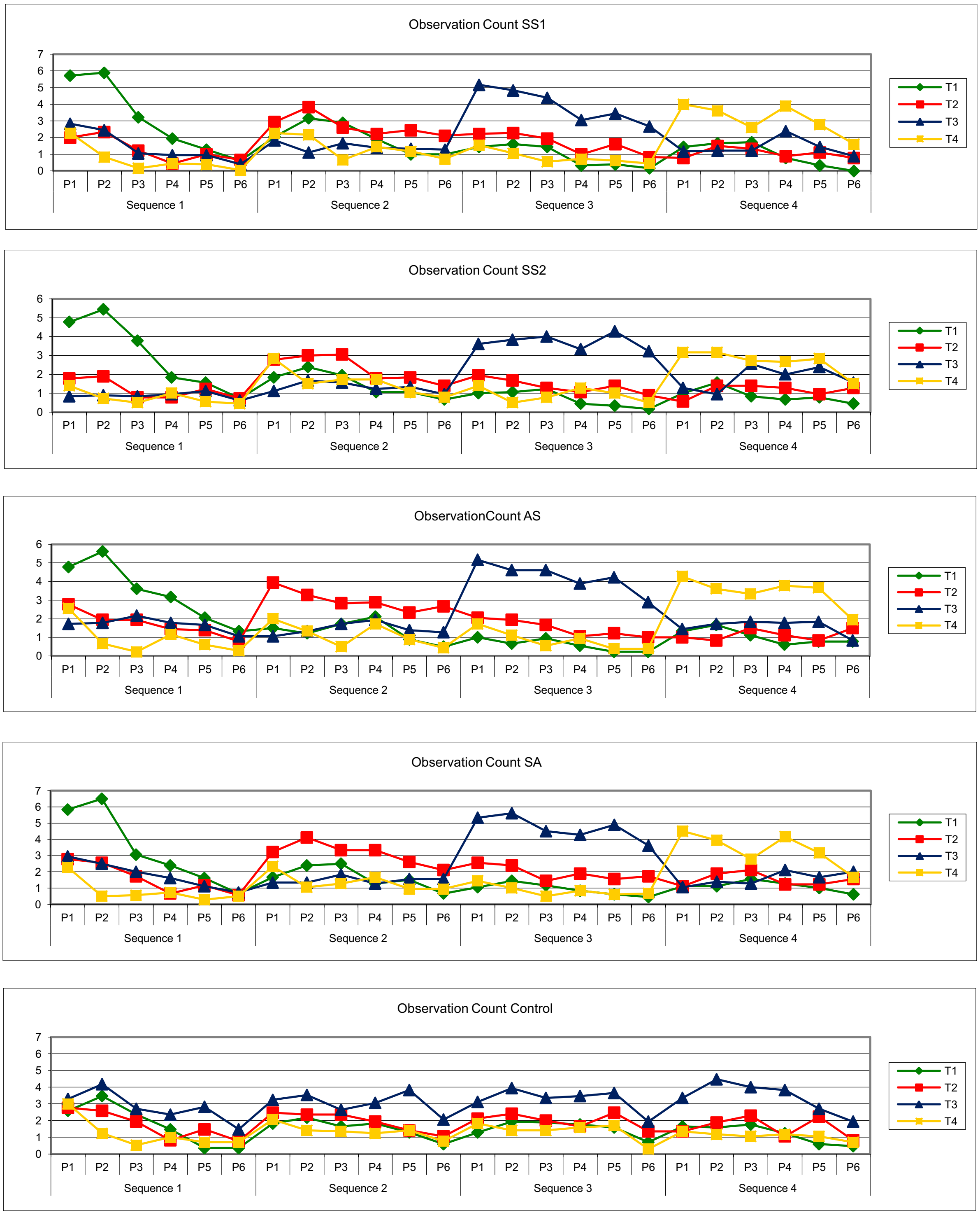


Figure 8.15. Mean observation count (number of visits) in each token (AOI), for each experimental condition (AS, SA, SS and AA and control); for each sequence (S), each picture (P1 to P6), and for each token (T1 toT4).

A repeated measured ANOVA was performed on the number of observations with presentation format (AS, SA, SS, AA and control) as between factors; sequences and tokens (AOIs) as within factors. This analysis showed a simple effect of presentation format conditions, $F (4, 84) = 5.1$, $p<0.001$, $\eta_p^2 = .20$. Post hoc analyses revealed that the number of observations was greater in the control conditions, compared with the AS, SA, SS and AA conditions. In these last condition learners spent more time, than in the control condition, on the cued tokens when the cueing appeared on the picture.

The effect of the sequence was significant, $F (3,252) = 3.71$, $p <0.02$, $\eta_p^2 = .04$. Post hoc analyses revealed that participants looked more the tokens sequence 3, and then followed by sequence 2 and sequence 1 and 4.

The effect of type AOI was also significant, $F (3,252) = 32.08$, $p<0.001$, $\eta_p^2 = .28$. Post hoc analyses indicated significantly that participants looked more at token 3, (AOI-3) (Damper and Hammer) than to others tokens.

More interestingly, the interaction between the type of token and the presentation formats modalities (AS, SA, SS, AA and control) was significant, $F (12,252) = 3.18$, $p<0.001$, $\eta_p^2 = .13$. This interaction revealed that participants in AS, SA, SS, AA conditions looked more often than in the control condition to the specific token cued at each sequence. This fact means that the number of observations in a specific token, once the token has appeared at a given sequence, was higher in the AS, SA, SS, AA than in the control condition, were the number of visits in the tokens were the merely same across sequences and for all tokens.

Result of eye tracking analysis for the animated presentation format for each condition (AS, SA, AA and control condition)

The following figures, 8.16 to 8.18, in below shows results of the average number of fixation before, total fixation duration, and observation count in each AOI, e.g. each token for each experimental condition for the animated format dual in each of the experimental conditions (AS, SA, AA1,AA2, Control).

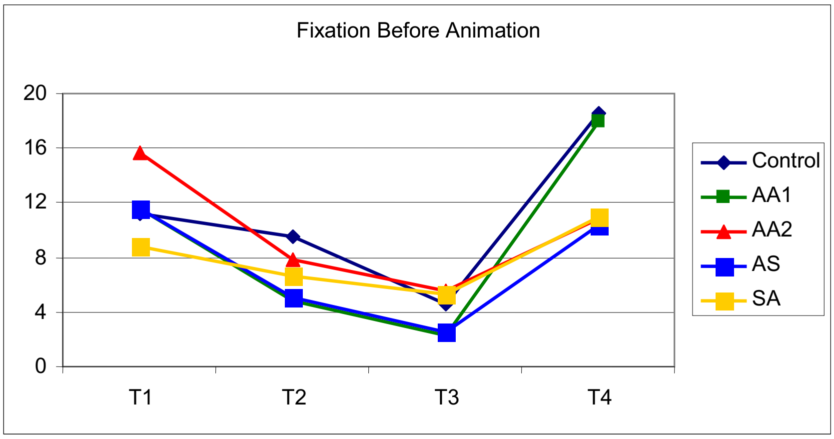


Figure 8.16. Mean number of fixation before the arrival in each token, for each experimental condition (AS, SA, AA1, AA2 and control) and for each token (T1 toT4)

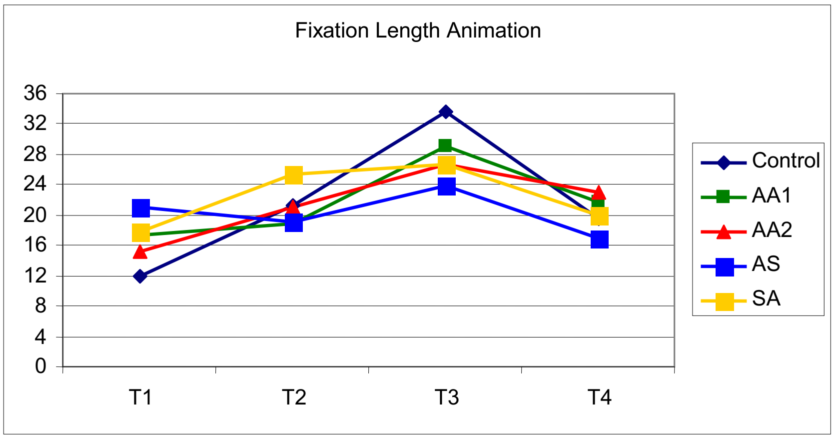


Figure 8.17.Mean total fixation duration in each token, for each experimental condition (AS, SA, AA1, AA2, and control) and for each token (T1 toT4)

Note: Token 1 (Key-Whippen), Token 2 (Hammer butt-Jack), Token 3 (Damper-hammer), Token 4 (Balance-Back check)

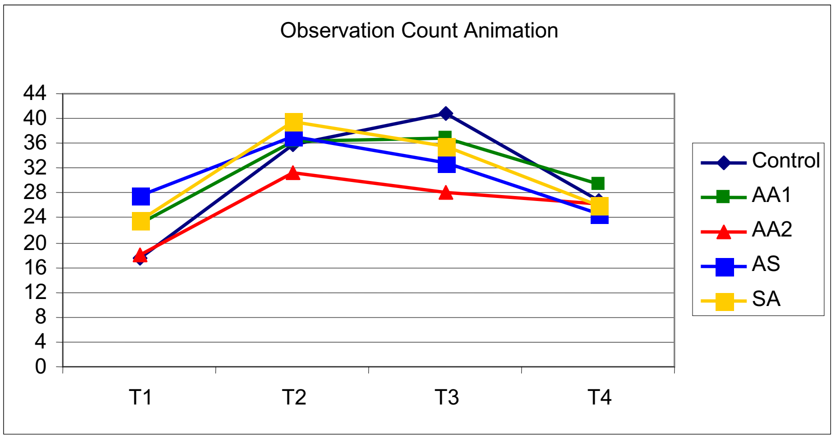


Figure 8.18. Mean number of observation in each token, for each experimental condition (AS, SA, AA1, AA2 and control) and for each token (T1 toT4)

For the number of fixation before the arrival in a token, a repeated measure ANOVA, with type of presentation (AS, SA, AA1, AA2, and control) as between factor and tokens as within factor was performed. Results showed no significant effect of presentation condition F (4, 84) = 1.59, p = 0.18. However, the analysis showed a significant effect of tokens, F (3,252) = 19.83, p<0.001, η_p^2 = .19; indicating that participants arrived sooner in the token 3. This token composed o the damper and the hammer have a great perceptual saliency, and the different experimental, conditions had no effect on this saliency effect.

For fixation duration, repeated measures ANOVA, a repeated measure ANOVA, using type of presentation (AS, SA, AA1, AA2, and control), as between factor and tokens as within factor was performed. Results showed no significant effect of presentation condition F (4, 84) = 1.77, p = 0.14. However, the analyses indicated a significant effect of the tokens, F (3,252) = 27.53, p<0.001, η_p^2 = .24, with the same meaning as the previous ANOVA, participants looked more at token three than at the others tokens.

For observation count, a repeated measure ANOVA, using presentation condition (AS, SA, AA1, AA2 and control) as between factor and tokens as within factor was performed. Results showed no significant effect of presentation condition F(4,84) = 2.19, p = 0.08. However, we found a significant effect of tokens types, F (3,252) = 53.70, p<0.001, η_p^2 = .39, in favor of tokens 2 and 3. The interaction between token types and presentation condition was significant, F(12,252) = 2.43, p<0.006, η_p^2 = .10.This interaction showed that participants in the AS and SA condition looked more at the more content relevant ad less salient tokens 1 and two than to the less content relevant and more salient tokens 3 and 4.

Executive function measures

In order to study possible links between the performance obtained in the comprehension post-test and executive function test, correlation measures were performed. A significant negative correlation was demonstrated between Post test 2 and errors in N-Back test ($r = -29$, $p < 0.007$).

Comprehension post test and Stroop test were positively correlated ($r = 0.23$, $p<0.03$) and with processing speed in the test comparisons of letter ($r = 0.21$, $p<0.05$).

A factorial ANOVA was performed on the performance of post test 2. The N back group (based on median split on the N back scores) and the type of presentation (AS, SA, SS, AA and control) were used as between factors. The analysis showed a significant effect of N-Back test group, $F(1,79) = 5.21$, $p<0.03$, $\eta_p^2 = .06$.

Conclusion

Contrary to experiment one, in experiment two, no effect of the dual format presentations was found compared to the single format presentations. This result did not seem due to the fact that we add cues in the visualizations. However cues directed efficiently learner's attention on tokens, and participants obeyed to the cues following in time the sequence of cued tokens. But seeing the right location at time on the relevant areas did not lead to efficient comprehension.

The analyses of eye movement recorded showed that participants observed more deeply the dynamic process in the dual format (SA and AS) than in the single format (SS1 and SS2). Participants tended to visit more frequently the AOI-3 (damper-hammer) with a sequential format viewing. Eye movement analysis indicated that participants observed more frequently the token 2 (hammer but-jack) and the token 3 (damper-hammer) that could be interpreted by the perceptual salience of those elements.

CHAPTER 9
EXPERIMENT 3. Cueing temporal features of series of static segmented presentation of dynamic processes followed by un-cued animation

Introduction

In experiment two, dual format presentations were compared to single format presentations (AS, SA, SS, AA). Temporal cueing on static pictures with tokens were used to foster inferences processes about the relation between the groups of components involved in the causal chain of the piano mechanism. However, in this experiment, cued tokens were used for the two formats: for the series of static pictures as well as for the animation. With such configuration of the cues applied on both formats, it was difficult, and quite impossible, to test precisely the potential effect of the cues in the static pictures on the processing of the following animation and on the comprehension test, or to test the effect of the cues in the static pictures on comprehension after the animation processing.

In order to overcome this difficulty, the third experiment presented in this chapter was the same as experiment two, except that there were no cues on the animation format presentation. In other word, temporal cueing of tokens was limited to the series of static pictures and the animation was no cued. Thus four conditions were designed and compared as follow: for the two dual formats, a series of static pictures, cued and an animation, no-cued; an animation, no-cued and a series of static pictures, cued; for the two single formats, two series of static pictures, cued; and two animations no-cued. So in this experiment no control condition was needed.

The same material as in the previous experiment was used and we assumed that the colour cues only in the static format will help to direct learners' attention to the relevant parts of the presentation and inhibit the information that seems inessential.

Research Goals

The goal of the experiment was to test the effect of dual versus single format presentation of the dynamic process of the piano mechanism. The effect of temporal cueing of tokens in the series of static pictures on comprehension was investigated.

The objective of experiment was determined whether cueing in learning piano mechanism system would result in better understanding than presenting the same format with un-cued animation.

The same as in previous experiment, eye tracking system was also used as a tool to get attention on perceptual processes in static sequential format presentation and animation format presentations.

Hypothesis

In this experiment, two formats of presentation, a series of static pictures showing the key steps of a dynamic process and an animation of the same process, are viewed in four conditions (the same as in experiment 2), however no cues were present in the animation: (1) Series of static pictures and an animation, **SA;** (2) An animation followed by the series of static pictures, **AS;** (3) a series of static pictures followed by the same series of static pictures, **SS** and (4) an animation followed by the same animation, **AA**. With regard to comprehension, it was predicted that participants in condition of **SA** and **AS** should get higher scores of comprehension scores than participants in both other conditions **AA** and **SS** (Hypothesis 1). Related to attention guidance, the absence of cues tokens in animation format presentation would influence in score because it was assumed that the rule of colour cues was considered to help direct participants attention to relevance parts of presentation (Hypothesis 2). For eye tracking measures, it was also predicted that the different conditions should demonstrate different patterns of attention according to the presentations formats.

Method

Participants

Forty eight students from Bourgogne University took a part as participants in the experiment. The design of experiment was shown in table 9.1. They were all over 18 years old and their average age was 26.7 years (SD=7.4) and 75% females. They were students of Psychology and they are given 0.5 credit point for their participation. Previously, participants have verified that they did not know exactly about how mechanism piano works. Students were assigned to four groups as table 9.1 shows.

Table 9.1. Design of experiment 3

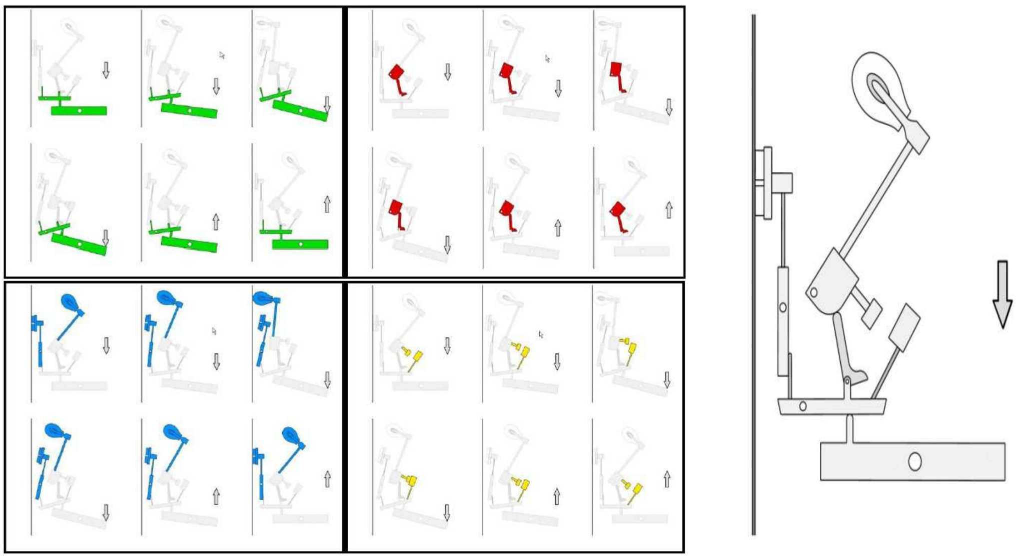
SECOND PRESENTATION	FIRST PRESENTATION	
	Static (S)	Un-cued Animation (A)
Static (S)	SS (N=12)	SA (N=12)
Un-cued Animation (A)	AS (N=12)	AA (N=12)

Material

In this experiment, we used the same as material in experiment 2, however the animation format was no-cued. All parts of the piano mechanism in the animation format were design with a neutral, but highly noticeable grey colour.

Experimental condition

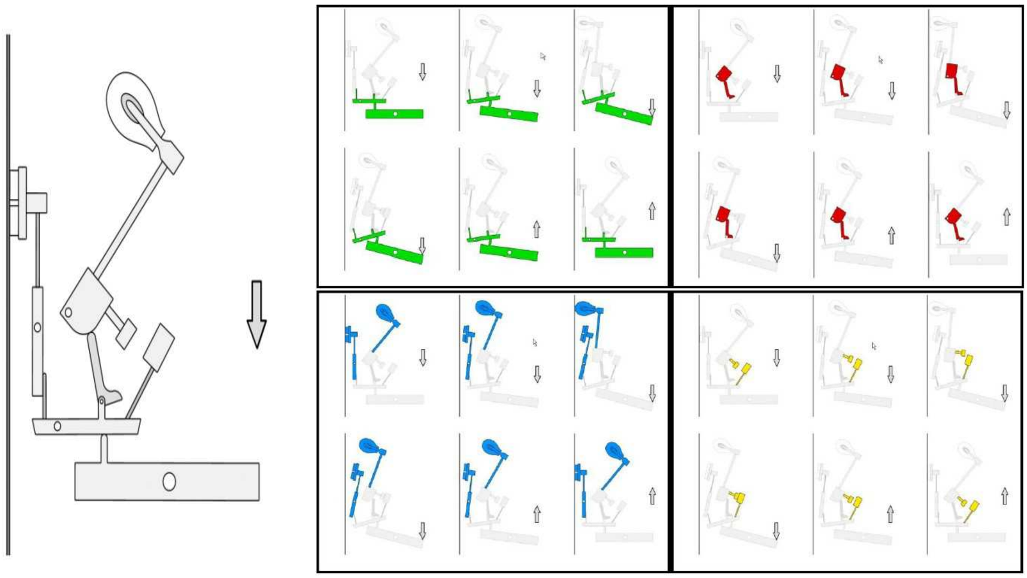
The experiment was designed in four experimental conditions: (1) a series of static pictures sequential with spatial cueing (tokens colored) and un-cued animation (SA- Figure 9.1), (2) an un-cued animation and a series of static pictures sequential with spatial cueing (AS-Figure 9.2), (3) a series of static pictures with cueing and the same series of static pictures with spatial cueing (SS- Figure 9.3), and (4) an no-cued animation and the same no-cued animation (AA- Figure 9.4).



1 minute 45 second

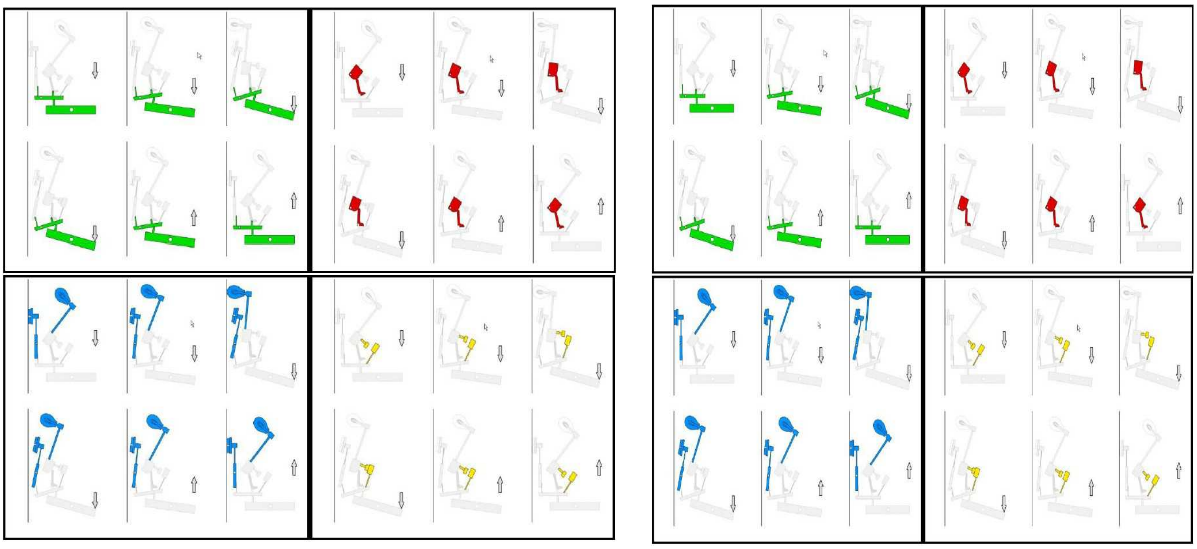
1 minute 45 second

Figure 9.1. Series of static picture followed by un-cued animation presentation (SA-condition)



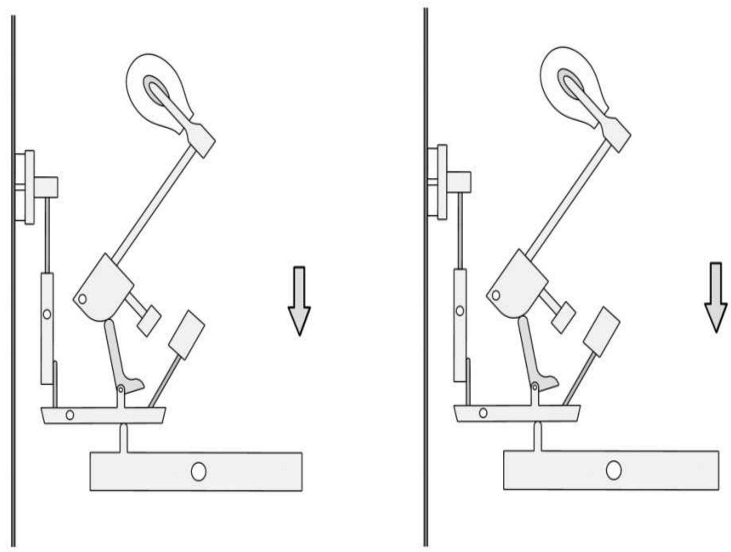
1 minute 45 second   1 minute 45 second

Figure 9.2. Un-cued animation presentation followed by series of static picture (AS-condition)



1 minute 45 second   1 minute 45 second

Figure 9.3. Series of static picture followed by series of static picture (SS-condition)



1 minute 45 second 1 minute 45 second

Figure 9.4. Un-cued animation followed by Un-cued animation (AA-condition)

The total duration of the animation presentation was also the same as in the previous experiment (one minute and 45 seconds). There was also a sound that happened when the hammer hit the string and visual signal or highlighting previewed when the hammer hit the string and when back catch contact to back check or the hammer rebounded the string. The parameters of temporal presentation were identical to experiment 2. Participants were also given Post test 1 (the local movement test) and Post test 2 (the functional mental model), and also the executive function tests.

Procedure

Participant sat in front of the screen at distance approximately 60 cm and eye tracking system started to calibrate eye movement for each participant before the experimentation started. The overall duration of each experimental condition was 3 minutes and 30 Seconds.

After completing the experiment, participant continued to carry out the post test (comprehension test, spatial ability test, and executive function tests). The total duration of each participant for doing our experiment was approximately one hour. Participant received credit point as reward to participate in this experiment.

Result

Comprehension test

The following table 9.2 showed the results of Mean in percent and (SD) for the first post test (the local movement test) and the second post test (the functional mental model).

Table 9.2. Means in percent and (SD) result of Post test 1 and Post test 2

Result	**Conditions**			
	SA (n=12)	AS (n=12)	AA (n=12)	SS (n=12)
Post test 1 (Local movement test)	47.36 (9.62)	45.76 (12.64)	47.01 (12.00)	47.92 (10.48)
Post test 2 (Functional Mental model)	31.11 (12.86)	29.17 (10.90)	31.39 (13.16)	26.39 (14.37)
Token 1 (Key-whippen)	48.61 (22.00)	20.83 (12.03)	47.22 (27.92)	40.28 (22.00)
Token 2 (H. butt-Jack)	17.71 (21.32)	14.58 (15.17)	20.83 (15.59)	20.83 (18.63)
Token 3 (Damper-hammer)	40.83 (13.20)	50 (15.81)	38.33 (17.24)	33.33 (19.72)
Token 4 (Balance-B.check)	15.28 (19.79)	22.22 (24.85)	18.06 (20.93)	8.33 (14.43)
Phase 1 (Striking)	26.85 (11.83)	29.17 (13.05)	33.33 (14.34)	25.46 (12.10)
Phase 2 (Rebounding)	9.79 (14.37)	20.83 (23.69)	18.06 (20.93)	4.17 (7.22)
Phase 3 (Resetting)	48.96 (23.08)	28.13 (16.24)	29.17 (19.98)	40.63 (26.08)

From calculated on all participants of the DAT4, the median score was 14.5 and for an overall average of 16.29. Participants were divided into two groups for the execution of statistical analysis. Subject with a score less than or equal to 14.5 was considered to have low spatial abilities (average score was 10.75) and those with scores higher than 14.5 was considered to have high spatial abilities (average score was 21.83). One-way ANOVA performed no significant difference between the mean score of DAT with the presentation conditions.

A repeated measure ANOVA, with presentation formats (AS, SA, SS, AA), DAT group as between factors and the post test 1 and 2 as within factors, showed no effect of experimental conditions, $F(3,40) = 0.12$, $p = 0.94$. However the analysis revealed a significant effect of spatial abilities on comprehension, participants with high spatial ability obtained higher scores than participants with low spatial ability (41.32% *vs.* 34.23% from mean corrected answer), $F(1, 40) = 4.69$, $p = 0.03$, $\eta_p^2 = .10$. Interaction between the type of post test and group level visual spatial was not significant, $F(1, 40) = 2.44$, $p<1$, ns.

A repeated measures ANOVA with presentation format condition (AS, SA, SS, AA), spatial abilities group as between factors, and tokens within factors showed a significant effect of token, F(3,120)= 23.94, p= .001, η_p^2 = .37. Post hoc analyses revealed that participants obtained significant higher scores for token 1 and 3 than compared with token 2 and 4.
There was also significant interaction between type of token and the presentation condition, F (9, 20)= 2.73, p = 0.006, η_p^2 = .17. Participants in the dual format presentation tended to perform better for the more relevant token for the comprehension of the mechanism (token 1 and token 2) than participants in the single presentation formats.

A similar analysis was performed on the scores for each phase of the post test. The analysis indicated a significant effect on the type of phase, F (2, 80) = 21.38, p < 0.001, η_p^2 = .35. Participants performed better phase 1 and 3 than the inconspicuous 2. Post hoc Fisher analyses showed that for scores of phase 3 were significantly higher than scores of phase 1 and score of phase 2. There was a significant interaction between the type of phase and the presentation condition, F (6, 80) = 2.88, p < 0.02, η_p^2 = .18. In the AA condition as well when the animation is presented before the series of static pictures, participants had a higher score for phase 2. This phase is more noticeable when a movement can be presented.

Eye movement

In this experiment, eye movement analysis was limited to the no-cued animation format which was the main focus here. Four AOIs were created with spatial boundaries and location which corresponded to the tokens (T1: key and whippen, T2: jack and hammer-butt, T3: hammer- hammer move and damper and T4: balance and back-check) cued in the series of static pictures. We still used three indicators: number of fixation before the arrival of the eye in the AOI, total fixation duration in each AOI, and observation count in each AOI. Result of AOI for fixation before, fixation length, and observation count in each experimental condition were showed in Figure 9.5, 9.6, and 9.7.

For the number of fixation before the arrival in each AOI, a repeated measured ANOVA was conducted with the experimental conditions (AS, SA, and AA with the two sequences of presentation AA1, AA2) as between factor and the AOIs (tokens T1, T2, T3, and T4) as within factor. The analysis showed a significant effect of experimental condition, F (3, 44) =

3, 39 p = 0.025, η_p^2 = .18. Post hoc Fisher analyses revealed that the number of fixation before the arrival of the eyes in the tokens, were lower in the SA condition than all other conditions. This result indicated that cued tokens presented before an animation was beneficial. There was also a significant effect of AOI (tokens), F (3, 132) = 17, 24, p = < .001, η_p^2 = .28. Post hoc analyses showed that AOI-1 and AOI-4 were viewed sooner higher than AOI-2 and AOI-3. The interaction between the effect of AOI and the experimental conditions was no significant, F (9,132) = 1.62, p = 0.11.

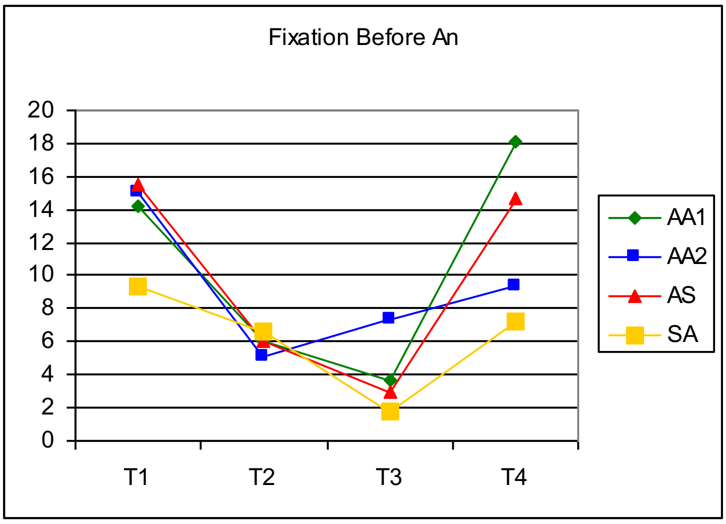


Figure 9.5. Mean number of fixation before the arrival of the eye in each AOI for each experimental condition (AS, SA, AA1, and AA2) and for each token type (T1 toT4)

A similar analysis for the total fixation duration in Figure 9.6 that used a repeated measured ANOVA with the experimental conditions (AS, SA, AA,) as between factors and the AOIs as within factor. The effect of the type AOI (tokens) was significant, F (3.132) =26.07, p = 001, η_p^2 = .37; showing that participants spent more time on tokens 2 and 3 than on tokens 1 and 4. Post hoc analyses revealed that AOI-3 had significantly longer fixation duration than all other AOI, while the AOI-1 had the shortest. There was also significant for interaction between the effect of type AOI and experimental condition, F (9,132) = 2.11, p = 0.032, η_p^2 = .13.

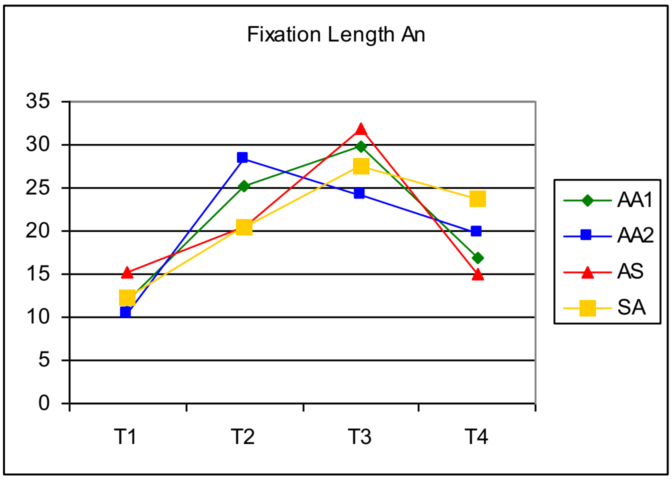


Figure 9.6. Mean number of fixation length for each experimental condition (AS, SA, AA1, and AA2) and for each token (T1 toT4)

For the number of visits in the AOIs (observation count, figure 9.7) a repeated measured ANOVA was performed with experimental conditions (AS, SA, AA1, and AA2) as between factor and the AOIs (four tokens) as within factor. The analysis showed a marginal effect of the presentation format condition, F (3, 44) = 2.43, p = 0.07, η_p^2 = .14. Participants in the AA2 condition had the lowest number of visits in the AOIs. The effect of the type of AOI was also significant, F (3,132) = 47.94, p= .001 η_p^2 = .52 in favor of AOI 2. However no significant interaction between the type of AOI and the experimental conditions, F (9,132) = 0.76, p=0.65 ns.

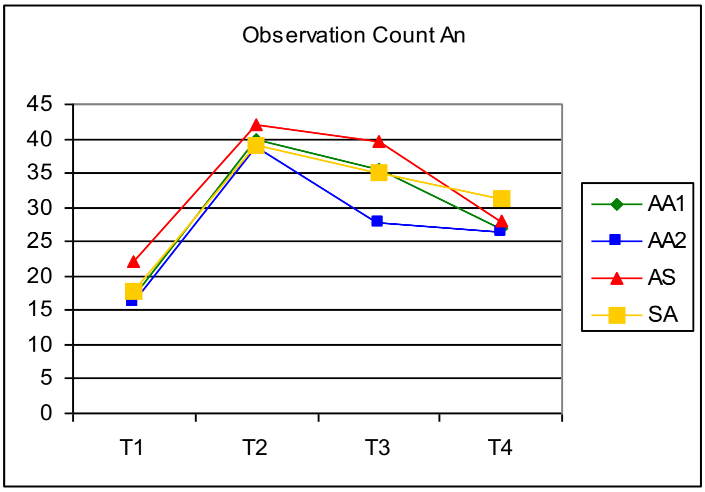


Figure 9.7. Mean number of visits (observation count) for each experimental condition (AS, SA, SS and AA) and for each token (T1 toT4)

Executive function measures

Correlation analyses were performed between the result of post test 1, post test 2 and the scores in the executive function tests.
No significant correlation was found between Post test 1 (and 2) and the number of errors in the N-back test ($r= -0.09$, $p = .054$).

Correlation analysis between the test 1 and the scores to the speed processing XO-comparison test showed a significant positive result, $r = 0.32$, $p < 0.03$. The same positive correlation was found between the Post test 1 and the Stroop test, $r =0.39$, $p < 0.007$.
An ANOVA with post test 1 as a dependent variable and Speed processing group (test XO) (based on the median split calculated on the XO scores) as between factor indicated a significant effect, $F (1, 40) = 4.81$, $p = 0.03$, $\eta_p^2 = .10$, in favor of the group with the higher scores.
Analyses of correlation for post test 2 and XO-test showed no significant effect, $F(1,40) = 1.37$, $p = 0.24$.

An ANOVA with post test 1 as a dependent variable, and Stroop test (interference task score) group (based on the median split calculated on the stoop test scores) as between factor also indicated a significant effect, $F(1,40) = 11.40$, $p < 0.002$, $\eta_p^2 = .22$. Similarly analyses for post test 2 showed a significant effect of Stroop test group, $F(1,40) = 5.26$, $p < 0.03$, $\eta_p^2 = .11$.

Conclusion

The present study investigated how visual cueing influenced attention allocation and cognitive processing when learning a complex mechanism presented with dual format, a series of static pictures cued and a no-cued-cued animation. We did found this time an overall significant effect of the dual formation presentation over the single formats, even using a no-cued animation. However and interestingly it seems that the best presentation format was the dual format with a series of static pictures depicting the key steps of the dynamic process followed by an animation.

Participants with high spatial ability obtained higher scores than participants with low spatial ability. Furthermore participants with high spatial abilities, who learnt in the dual format with cueing condition performed better than the high spatial participants in the single format presentation.

CHAPTER 10

EXPERIMENT 4. Comparison of the effect of 2D versus 3D animation presentation on the comprehension of a complex animation, with an eye tracking investigation

Introduction

Two D and 3D animations in research has been used for a long time in areas such as military environment (Smallman, John, Oonk, and Cowen, 2000), air traffic controllers (Smallman. St. John, Oonk, and Cowen, 2001), geographic terrain (Hollands Justin G and Ivanovic Nada, 2002), topographic maps (Wiebe, Slykhuis, Madison, & Savage, 2004) and medical areas (Agus, Fabio, Enrico, & Giovany, 2007) but not in the educational area.

In the line of the work by Lowe and Boucheix, 2010 and Boucheix, & Lowe, 2010; in this research, the difference between 2D and 3D presentation of the animation is used as a way to manipulate the perceptibility profile of the presentation of the mechanical systems that is to say to manipulate the “natural” cueing effect, here the “natural” visual saliency of the components of the piano mechanism. The level of perceptibility of each component is different in 3D compared to 2D. For example, in 2D the salient (but less content relevant) hammer has a high level of perceptibility and the no-salient (but highly content relevant) whippen has a low level of perceptibility. In 3D, because of the depth feature, the whippen could now acquire a higher level of perceptibility. Thus the salience relevance misalignments seem weaker in 3D than in 2D as shown in figure 10.1.

As in the other experiments, the chosen system, the piano mechanism, was used in the research because of people’s general unawareness of its functioning, despite of its fairly complex, but still clear, sequence of relevant events. And, most interesting there was little correspondence between the perceptual salience of the mechanism’s individual components and their thematic relevance to its function (see, Lowe & Boucheix, 2008; Boucheix and Lowe, 2010; Lowe & Boucheix, 2011).

We predicted that 3D more than 2D animation will affect attention and comprehension of learners. We used also an eye tracking system because it provides mainly information on visual attention direction and, in this research; eye tracking indicators were employed on the basis of dynamic area of interest (AOI) created with new software. It was also predicted that the eye tracking measures would show different patterns of attention direction relative to specific parts of piano mechanism.

In the experiment presented below, with the 2D versus 3D factor, a second factor was manipulated: the type of focus on the components of the animated piano mechanism presented to the learner during the study of the animation. Three system paced modalities of focus on the animation were used and presented to the learner during the operation of the piano mechanism (i) a visualization focused on the whole animation (all components visible) at the same time during all the study time, (ii) an alternate visualization of the whole piano mechanism and of zoom in and zoom out focus on different random parts (random group of components) of the piano mechanism (iii) an alternate focus on the whole mechanism and on specific content relevant related functional components (see figure 10.3 to 10.5).

It was expected that learners in the alternate focus on the whole mechanism and on specific content relevant related functional components would have higher comprehension performance than learners in the alternate visualization of the whole piano mechanism with zoom in and zoom out focused on different random parts of the piano mechanism. The lowest performances were predicted for the whole animation at the same time focus only.

Research Goals

In this research, the goal was mainly to compare 2D and 3D animations for the understanding of mechanical systems. Eye movements were recorded during the presentation of 2D and 3D animation of piano mechanism system. Is there any difference in eye fixations and performance between 2D and 3D presentation of a complex mechanical system?

We were particularly interested in the study of the difference between 2D and 3D animation in term of perceptibility profile of the piano animation. Eye movement indicators were used to answer these questions. The number of fixation in different area of interest (AOI) of the piano (key, whippen, jack, hammer-butt, hammer, damper spoon, damper, balance back check, and string) was an eye data indicator that could bring information about learner's attention and perceptual processes. In the experiment was used "ocnoscere", a new software allowing to perform eye tracking analysis on dynamic areas of interest (AOIs) (Casallas, 2009). This was the first experiment in which this software was used for the analysis with dynamic area of

interest. Previous research used mainly analysis with static area of interest (AOI). In this experiment, we focused on different part for 2D-3D animation in three visualization focus groups: whole animation, animation with zoom in-out alternatively on the whole components and on random groups of components, and animation with zoom-in-out alternatively on the whole components and on specific content relevant related groups of functional components.
The comprehension test was inspired by Narayanan and Hegarty (1998), and concerned (a) the configuration of the mechanical system, (b) the local kinematics and (c) the functional mental model of the piano system (Narayanan & Hegarty,M, 2008).

METHOD

Participants

Ninety undergraduate students at University of Burgundy in France participated in the study with a mean age of 18.7 years (SD = 1.16); they participated for course credit. They were divided into 6 independent groups; each group consists of 15 subjects. The design of experiment is shown in table 10.1.
The experimental procedure followed three phases: (i) a pre-test with questionnaire about prior knowledge about piano mechanisms, (ii) a study time of the piano animation with the recording of eye movement by an eye tracker, and comprehension post-test. The system that was used had no user control and the same time constraint (3’30”) for each subject.

Table 10.1. Design of experiment 4

Focus group	**Type of presentation**	
	2D	**3D**
the whole piano (system paced)	N=15	N=15
Zoom in-zoom out on each different parts(system paced)	N=15	N=15
focused in specific relevant related functional (system paced)	N=15	N=15

Material
2D and 3D animation presentation

Piano mechanism is showed in 2D and 3D version animation without text accompanying the animation (Figure 10.1 and Figure 10.2). Three versions focus group of animation piano mechanism were presented in each two and three dimension: (a) the whole piano mechanism: animation showing the whole mechanism without focusing on any of the parts of piano (b) zoom in-out on each different parts of piano: animation showing the whole mechanism then focusing on AOIs by zooming in and out on each different parts of piano and (c) focused in related between parts of piano: animation showing the whole mechanism then focusing on AOIs that representing the relations between the different parts of the piano with zooming in and out on each of them. A part of three versions focus groups of animation is captured by viewed in figure 10.3, figure 10.4, and figure 10.5. "Ocnoscere" was used in eye tracking analysis for dynamic area of interest. An example of the output of the "ocnoscere" is presented in figure 10.6. The "ocnoscere" software was created in our laboratory, and used especially for Tobii eye trackers.

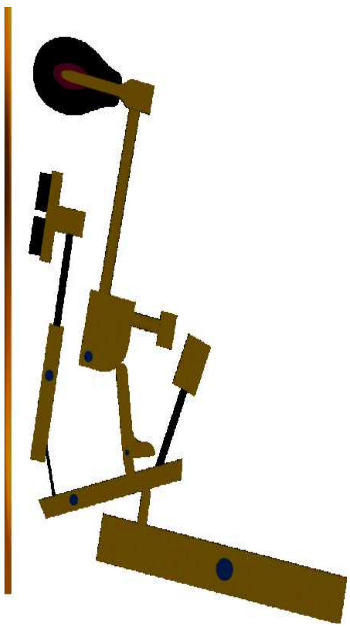


Figure 10.1. 2D version

Figure 10.2. 3D version

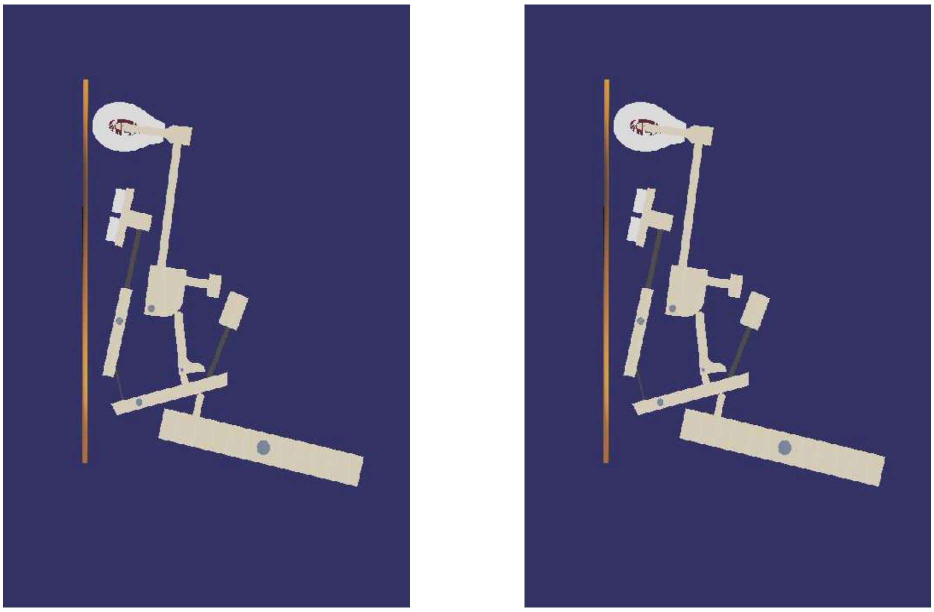


Figure 10.3 A visualization focused on the whole animation (all components visible) at the same time during all the study time

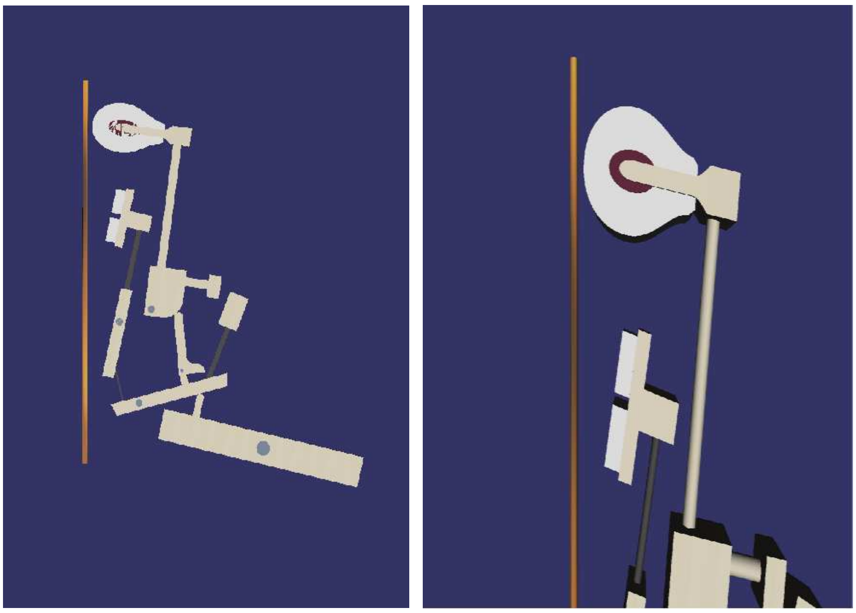


Figure 10.4. An alternate visualization of the whole piano mechanism and of zoom in and zoom out focus on different random parts (random group of components) of the piano mechanism

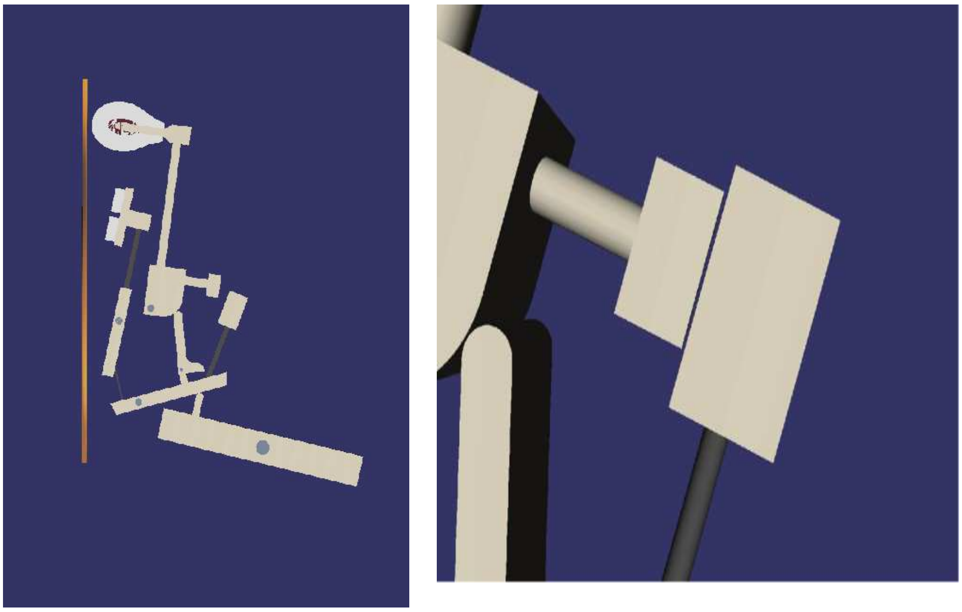


Figure 10.5. An alternate focus on the whole mechanism and on specific content relevant related functional components

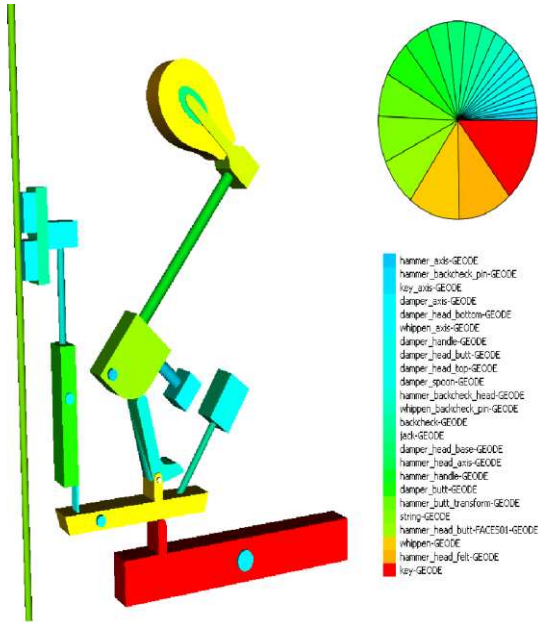


Figure 10.6. An example ocnoscere output, here each component with a specific color is an AOI which move. Each color represents a dynamic are of interest. The pie chart beside the piano mecahnism shows the number of gazes of a participant in each of the dynamic AOIs.

Comprehension test

Comprehension test was a written questionnaire given to each learner after the learning time items concerned (a) the configuration of the piano system's parts, (b) the local kinematics of the system and (c) the functional mental model of the piano system.

The configuration of the piano system's parts and the local kinematics of the system were in the questionnaire format (Table 10.2) and the functional mental model of the piano system was in the written format or essay test. Participants had to write precisely the following question:

"What happens with all parts of the piano when a pianist presses the key down and then releases it". Previously we have given a sheet that shows, in a random order, the elements of the piano with their names (see figure 7.8). Scoring criteria for functional mental model test is the same test as the scoring grid which was used in experiment 1 (see on table 7.2). Scoring criteria was based on 15 micro-steps constituting the three main stages of a piano mechanism's functioning (Boucheix and Lowe, 2010).

Eye tracking equipment

Eye movements were recorded with the Tobii 1750 video eye tracker at a constant 60Hz frequency. The screen was positioned approximately 60cm away from the participant. The fixation filtering threshold was set at 20 *ms*. Ocnoscere software was used for the eye movement analysis.

Table 10.2.Comprehension test: the configuration of the piano system's parts, the local Kinematic of the system piano mechanism, and the functional mental model.

Name:			
Age:			
The configuration of the piano system's parts and the local kinematics of the system piano mechanism			
Please answer these questions below by the check list the correct answer			
When the pianist presses on the key:			
1.	The whippen pushes the damper to the string	TRUE	FALSE
2.	The whippen is in contact with the damper spoon	TRUE	FALSE
3.	The hammer rebounds on the string	TRUE	FALSE
4.	The key is in contact with the hammer's butt	TRUE	FALSE
5.	The damper causes the hammer to hit the string	TRUE	FALSE
6.	The damper spoon is in contact with the key	TRUE	FALSE
7.	The balance back check is in contact with the hammer butt	TRUE	FALSE
8.	The damper moves away from the string	TRUE	FALSE
9.	The end of the key pushes on the whippen's right	TRUE	FALSE
10.	The hammer butt is in contact with the hammer	TRUE	FALSE
11.	The jack causes the hammer to hit the string	TRUE	FALSE
12.	The jack's position is fixed under the hammer butt	TRUE	FALSE
13.	The balance back check moves on the top of the whippen	TRUE	FALSE
14.	The damper spoon carries on the withdrawal of the damper from the string	TRUE	FALSE
15.	The hammer butt is in contact with the whippen	TRUE	FALSE
16	The damper spoon moves on (top of) the whippen	TRUE	FALSE
17.	The whippen does not rotate	TRUE	FALSE
18.	The whippen carries on the jack to push the hammer butt	TRUE	FALSE
19.	The whippen carries on the withdrawal of the hammer pushing of the damper	TRUE	FALSE
20.	The whippen's movement carries on the hammer butt which causes the hammer butt hit the string	TRUE	FALSE
21.	The balance back check is pushed by the back check	TRUE	FALSE
22.	The whippen pushes the back check towards the hammer balance	TRUE	FALSE
23.	The key is in contact with the whippen	TRUE	FALSE
The Functional mental model test			
24. What happens with all parts of the piano when a pianist presses the key down and then releases it? Please explain it as precisely as possible.			

Eye movement data analysis

We created nine areas of interest (AOIs) which were used to record the activity of eye movement's subjects when 2D and 3D animation were presented. Eye movements of participants were analyzed during experiment in each experimental condition, as shown in Figure 10.3, 10.4, and 10.5. We defined 9 dynamic Area Of Interest: AOI-1 (Key), AOI-2 (whippen), AOI-3 (Jack), AOI-4 (Hammer-Butt), AOI-5 (Hammer), AOI-6 (Damper spoon), AOI-7 (Damper), AOI-8 (Balance-Back check), and AOI-9(String) (see Figure 10.7).

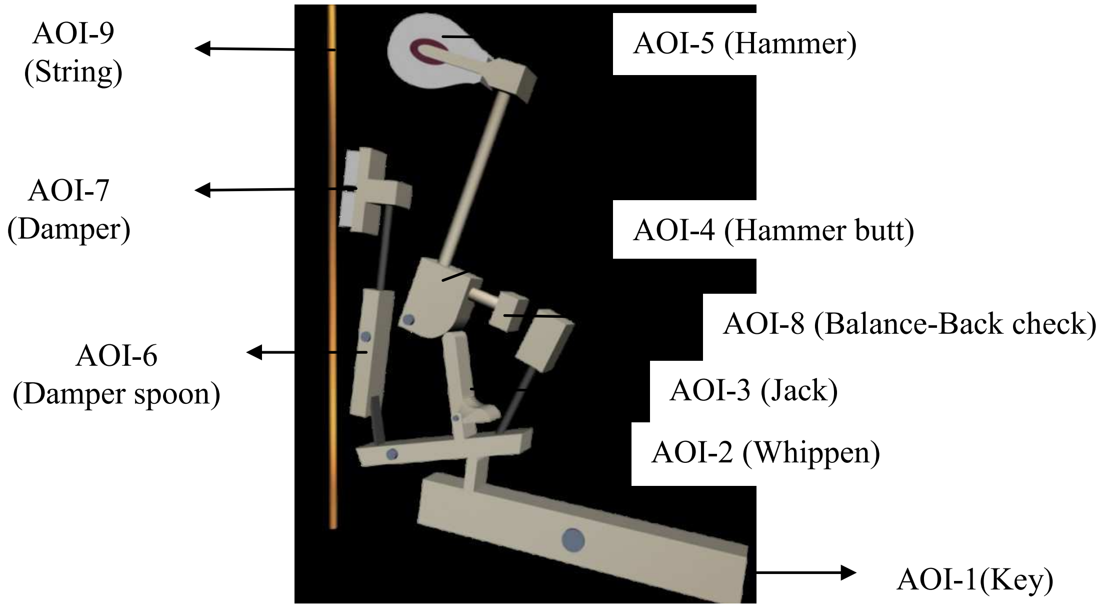


Figure 10.7.Nine dynamic Area of Interests (AOIs) of the piano mechanism system.

Procedure

Subjects were first given a pre-test to evaluate their knowledge about the mechanical piano system. After answering the test, the eye tracker was calibrated for each participant. Eye movement were recorded by o*cnoscere* software and the study was limited to three minutes. After having completed their study of the animation, subjects solved the comprehension test.

Results

For this experiment, we were mainly interested in the problem of the perceptibility profile of the animation and we choose deliberately to present first the result of the eye tracking measures and to present secondly the comprehension measures.

Eye tracking measures

We analyzed the eye gazes of ninety subjects, table 10.3 and figure. 10.8.

A repeated measures ANOVA on the number of gazes in each of the different AOIs was performed with 2D and 3D animation group and type of focus group as between factors and AOIs as the within factor. The analysis showed a significant effect of 2D vs. 3D Group, F(1,78) = 16.3724, $p = 0.000122$ and a significant effect of the type of group focus, F(2,78) = 4.5512, $p = 0.013506$.

Analysis showed that there was significant interaction between the number of gazes in each AOIs and group 2D and 3D, F (8,624) = 8.9291, $p = 0.00001$. For a group of AOI (key, whippen, Jack, hammer, damper and balance back check), eye gazes number was higher in 3D than in 2D, (Figure. 10.8).

Table 10.3. Number of gazes and SD for each AOIs in group 2D-3D

	Key	Whippen	Jack	Hammer Butt	Hammer	Damper Spoon	Damper	Balance Back check	String
2D	317.98 (238.10)	605.67 (369.23)	476.64 (276.61)	626.64 (417.90)	678.40 (536.15)	39.07 (68.85)	582.84 (302.28)	536.84 (340.91)	59.16 (58.85)
3D	541.51 (363.75)	703.67 (346.67)	661.40 (357.16)	645.09 (365.44)	1371.53 (1251.84)	88.69 (110.60)	685.31 (333.96)	737.71 (374.02)	58.98 (67.01)

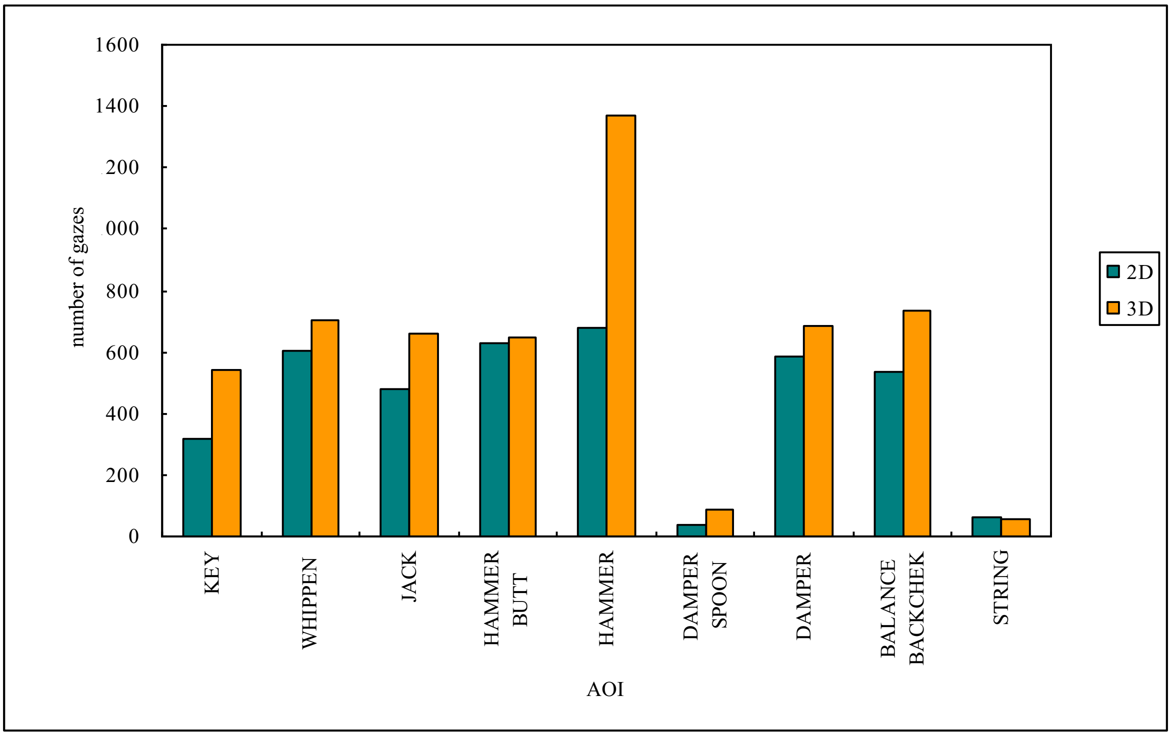


Figure 10.8. Number of gazes and AOIs in group 2D-3D

ANOVA showed also significant relationship between the number of gazes in the AOIs and the type of focus group F, (16,624) = 12.655, p = 0.0000. The mean number of gazes and standard deviation (SD) for each AOI, with the detail for each type of focus group is presented in table 10.4 and figure 10.9.

Table 10.4. Mean number of gazes and SD in each AOIs for each group focus

	Key	**Whippen**	**Jack**	**Hammer-Butt**	**Hammer**	**Damper Spoon**	**Damper**	**Balance-Back check**	**String**
Whole piano	359 .9 (268.82)	464.9 (297.27)	468.2 (384.64)	589 (417.38)	819.4 (682.30)	17.2 (43.79)	666.16 (365.03)	502.9 (383.20)	53.2 (58.09)
Zoom in-Zoom out	537.67 (409.14)	603.7 (364.60)	516.6 (285.39)	730.33 (462.23)	1614.4 (1453.30)	21.93 (24.84)	550.03 (320.45)	680.06 (444.07)	27.86 (32.4)
Focus in related part	391.56 (259.30)	895.4 (274.65)	722.4 (262.70)	588.26 (257.16)	641.1 (231.07)	152.5 (113.69)	686.03 (262.48)	728.86 (215.83)	96.13 (71.96)

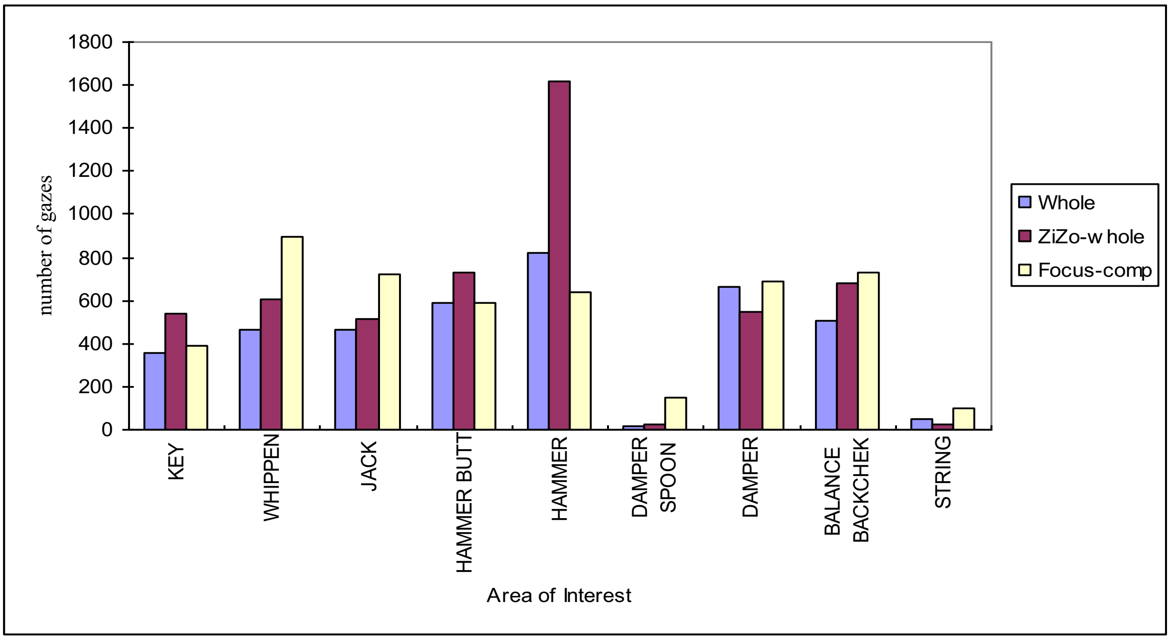


Figure 10.9. Relation between number of gazes in each AOIs with the type of focus group.

There was a marginal effect for the interaction between the type of focus group and the 2D/3D factor, $F (2, 78) = 3{,}0346$, $p = 0.05381$ (table 10.5 and figure 10.10).

Table 10.5. Total number of gazes in group focus and in group 2D-3D and SD

Group	Whole piano	Zoom in- Zoom out	Focus in related part
2D	395.28 (446.70)	420.39 (382.46)	525.44 (445.05)
3D	480.45 (434.27)	753.51 (914.04)	597.32 (332.62)

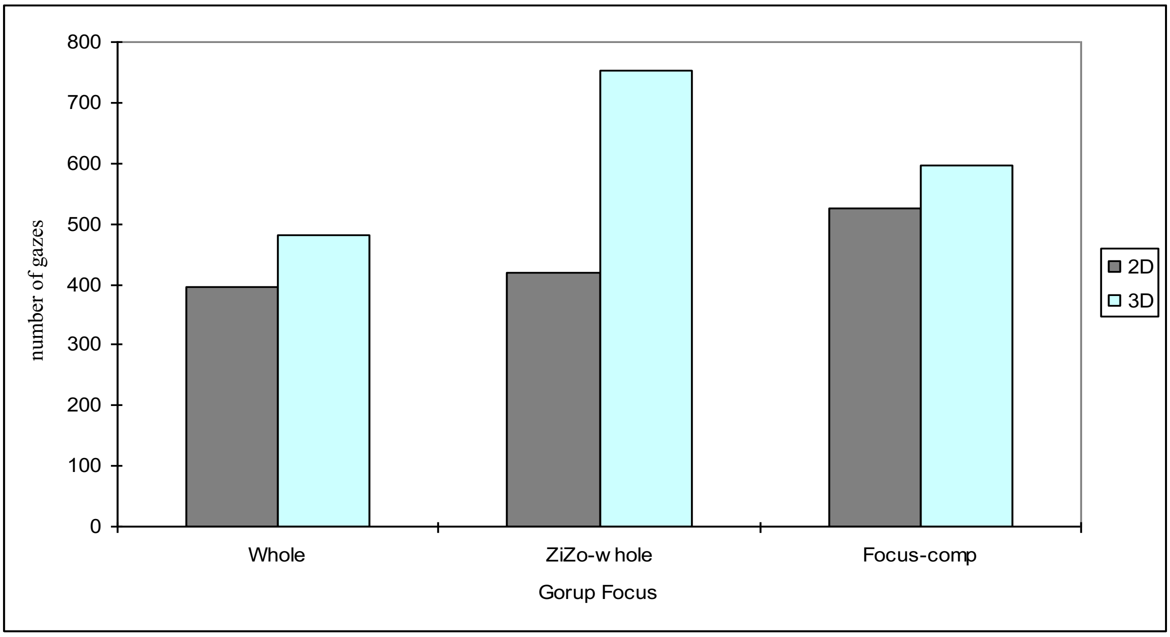


Figure 10.10. Interaction between group focus and group 2D-3D

The number of gazes in the 3D group was higher for the group zoom in-out with random groups of components than for the group zoom in-out with relevant content related functional parts of the piano. These two last groups had also more gazes that whole components group. In group 2D, the number of gazes in the group focus with relevant content related functional parts of the piano was the highest. This result showed that the participants spent more time and were more focused when animation showed zoom in-out alternation for 3D animation. A last analysis showed a significant double interaction between, the 2D-3D group, and type of focus group focus, F (16,624) = 15.2380, $p = 0.0000$.

Comprehension measures

Figure 10.11 and figure 10.12 show the comprehension test scores in percentages, respectively as a function of the type of focus group and as a function of the 2D vs. 3D group.

Table 10.6 and Figure 10.11 show percentages of comprehension measures. The total of comprehension measures in group 2D and 3D can be showed in table 10.7 and figure 10.12.

Table 10.6 Percentages of total measures comprehension and SD

Group Focus	Group	% Total measure comprehension and SD	Mean (SD)
Whole	2D	52.71 (11.96)	54.47 (2.48)
	3D	56.22 (13.42)	
Zoom in Zoom out	2D	48.59 (12.90)	50.44 (2.61)
	3D	52.28 (11.86)	
Focus in related	2D	49.12 (15.81)	50.12 (1.43)
	3D	51.14 (12.71)	

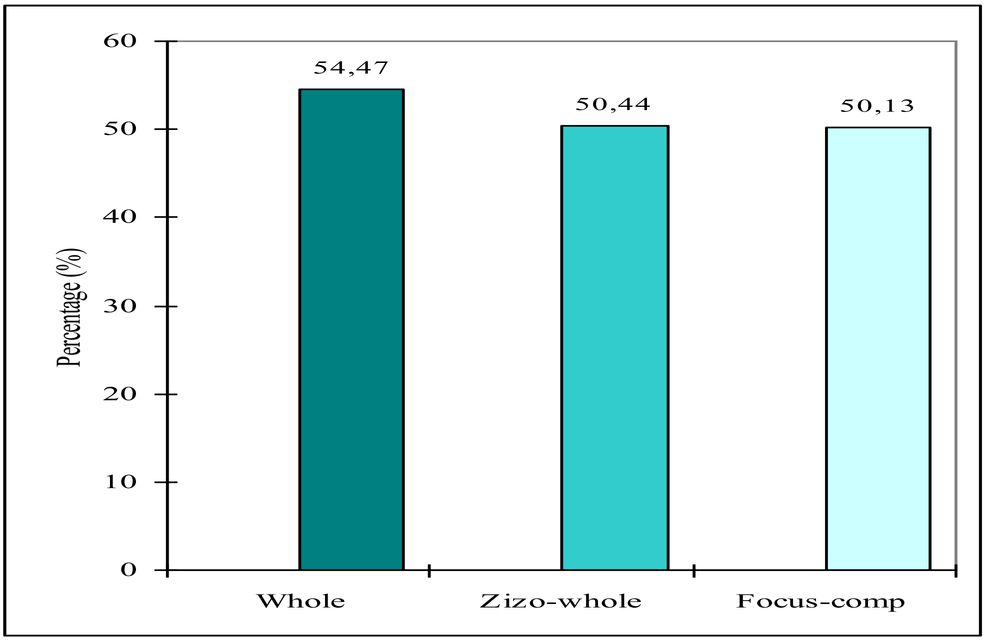


Figure 10.11.Percentages of comprehension measures in the different type of focus group.

Tabel 10.7 Percentages of total comprehension measures in group 2D and 3D

Group	% total comprehension measures and SD
2D	50.14 (2.24)
3D	53.21 (2.66)

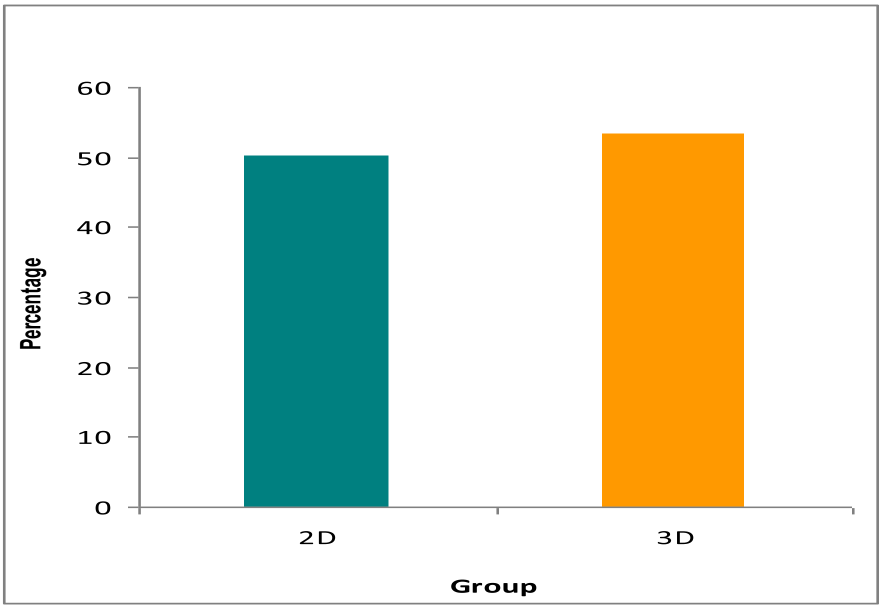


Figure 10.12. Total comprehension measures in 2D and 3D groups

A repeated measure ANOVA on comprehension scores with 2D and 3D group, type of focus group as between factor and comprehension sub scores (configuration, local kinematics, and functional mental model) as within factor was performed. This analysis showed a significant effect 2D vs. 3D group in favour of the 3D presentation, $F(1, 78) = 5.814$, $p = 0.018$, but no significant effect of the type of focus group.

Discussion

The result of this experiment showed that the use of 2D and 3D presentation could change the perceptibility profile of an animation. This modification could result in attention direction and comprehension scores differences. As in Hegarty (2011) stated recently what's matter in learning with visual displays is what you see (and not how you interact with).

Of course, the eye tracking measures showed that the number of gazes was higher in 3D than in 2D animation. The means number of gazes on hammer and whippen both in 2D and 3D animation were the highest. In fact, hammer and whippen are crucial component of the mechanism, the hammer has is very salient while whippen is less salient and more relevant. It might be that 3D presentation result in more eye fixations in numbers of components of the display. The results are consistent with those found by Boucheix and Lowe (2010) regarding the perceptibility profile of the piano mechanism animation. Because the hammer is a large shaped item whose movement rapidly covers much of the display area, so it has a high level of perceptibility.

The results are similar with those held by Pavclinic, Paul, Janet, and Tony (2001) but in a different field. They compared 2D and 3D animations in understanding stereochemistry and found that 3D animations and colourful molecular images encouraged students to practise and understand better Stereochemistry.

Conclusion

The main conclusion is that 3D animation was better for directing learner’s attention and for improving comprehension than 2D animation, even if in this particular case depth feature was not needed to understand the causal chain of the mechanism.

CHAPTER 11
SUMMARY OF RESEARCH AND GENERAL DISCUSSION

11.1. Summary of the research and comparison between experiments

The purpose of our research were (1) to examine whether the building of mental models from animation could be improved by using external simplified model representations of the dynamic reference content (Experiment 1), (2) to study the cognitive process on learning complex system with cueing as direction of visual attention and eye tracking system as a tool for "on line" measures on recording of perceptual processes in complex static and dynamic format presentations (Experiment 2), (3) to observed the effect of absence cues token for an animation presentation in learning piano mechanism system (Experiment 3) and (4) to learn 2D and 3D animations for the understanding of learning complex animation (Experiment 4).

Hypothesis of our research are (1) dual format representation is better than single format representation– Experiment 1; (2) Effect of cueing guidance for building mental model in all experiment condition– Experiment 2; (3) Cueing increased score in dual format than single format– Experiment 2; (4) The absence of cued tokens in the second format presentation would influence in score, because it was assumed that the rule of colour cues was considered to help direct participants attention to the relevance parts of presentation – Experiment 3, (5) For an eye tracking analysis, static series pictures format with colours is expected to generate fixation times greater for tokens than animation format without tokens– Experiment 3; and (6) 3D animation will affect attention and comprehension of learners for depth than 2D animation – Experiment 4.

In this chapter we will present a summary of the main results of our studies then we will expose the limits of our experiments. Finally, we have some perspectives of research before making ergonomic information to design more effective format presentation in learning from a complex animation. Table 11.1 shows the summary of conditions of our studies with the columns that specify the experiments, treatment, objectives, hypothesis, the main results, and conclusions for each experiment.

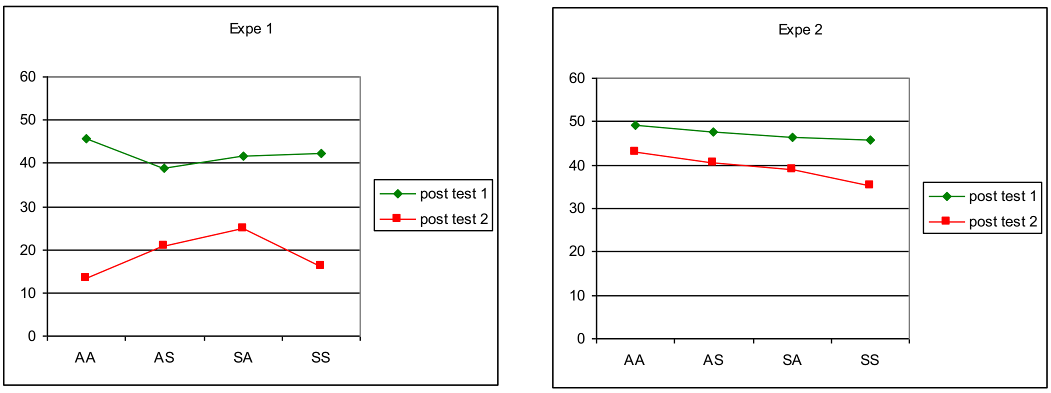
Table 11.1. Summary of the conditions of our studies

Experiment	Treatment	Objectives	Hypothesis	Results	Conclusion
Experiment 1: Chaining two types of formats in order to foster inferences for building high quality mental models	Combined a series of static pictures and an animation in different position.	To examine the effect of the presentation of simplified external representations on learning from a complex animation.	An integrated static sequential presentation and an animated presentation should be better for developing a dynamic mental model than single format presentation.	SA and AS format is better than AA and SS format	Dual format presentation is better for developing a dynamic mental model than single format
Experiment 2: The use of visual 'tokens' in cueing temporal features of series of static pictures in dynamic process presentation.	Cueing spatial temporal "tokens" of series of static pictures with a cued animation. -Using an eye tracking analysis	• To investigate whether strategies of dual format with cueing would result in better comprehension than single format presentations and eye tracking system as a tool for "on line" measures. • To investigate the effect of executive functions on understanding the complex mechanism system	• Strategies of dual format with cueing would result in better comprehension than used single format presentations. • Positive correlation between the level of executive functions and performance in comprehension of the complex mechanism system	• Participants observed more in SA and AS than SS1 and SS2. • When animation was presented, participants tend to look frequently to AOI-2 and AOI-3. • A significant negative correlation was demonstrated between Post test 2 and N-Back test	• Cueing can guidance for building mental model and enhance performance in dual format presentation. • Eye movement measures stated that the perceptual salience of hammer-damper, and hammer-but and jack when viewing the animation.
Experiment 3: Cueing temporal features of series of static segmented presentation of dynamic processes followed by un-cued animation	Cueing spatial temporal "tokens" of series of static pictures with un-cued animation. -Using an eye tracking analysis	• To determine whether cueing would result in better understanding than presenting the same format with animation without cueing. • To observe the effect of absence cues token for an animation presentation.	• The absence of visual cueing on animated format presentation should affect perception and interpret cognitive process. • Static series pictures format with colours is expected to generate fixation times greater for tokens than animation format without tokens	• Participant observed more in AOI-2 and AOI-3. • Participants look more often in SA and AS condition than in AA1 and AA2 condition.	• Participants looked more often and longer periods of time at visual cueing in static format presentation than at non visual cueing in animated format presentation.
Experiment 4: Comparison of 2D/3D presentation with an eye tracking analysis	- 2D/3D of perceptibility profile in depth. - Interactivity: whole animation, zoom-in/out, focused on related part.	To examine the difference between 2D and 3D of perceptibility profile in piano animation.	3D animation will affect attention and comprehension learners for depth than 2D	3D was better than 2D	3D animation was better in affect attention and comprehension for depth than 2D animation.

Comparison of experiment 1 and 2

The studies presented in this dissertation were set up with the same system of piano mechanism and we investigated whether attention cueing with token and temporal cueing are an effective means for enhancing learning from complex animation. Thus we can compare the results of each experiment.

The local movement scores (Post test 1) and the functional mental model scores (Post test 2) of the two experiments are compared (Figure 11.1). We used ANCOVA analyses including the group level of DAT as continuous factor and experimental conditions as categorical factors. The analysis showed significant effect for post test-1, F (1,145) = 8.09, $p < 0.006$, $\eta_p^2 = 0.05$, but no significant effect for interaction with the type of experiment F(3,145)<1, ns. Similar with Post test-2 showed significant F (1,145) = 92.73, $p<0.001$, $\eta_p^2 = 0.05$. Interaction between type of presentation is significant, F (3,145) = 2.92, $p<0.04$, $\eta_p^2 = 0.06$. Post hoc test of Fisher indicated that average difference between two experiments was significant for all conditions of presentation. Comparison between two experiment showed that performance of subjects on condition AA increased significantly more than in other conditions, $p<0.006$.



(a) (b)

Figure 11.1. The result of comparison for post test 1 and post test 2 between (a) experiment 1 and (b) experiment 2.

Comparison of experiment 1 and 3

A similar analysis was conducted to compare the result of experiment 1 and experiment 3, showed that score the local movement (Post test 1) for experiment 3 significantly higher than score in experiment 1, F (1,121) = 4.85, $p<0.03$, $\eta_p^2 = 0.03$.

The same as for the functional mental model (Post test 2), score in experiment 3 showed higher than experiment 1, F (1,121) = 22.27, p<0.001, η_p^2 = 0.15 and no interaction with all experimental conditions.

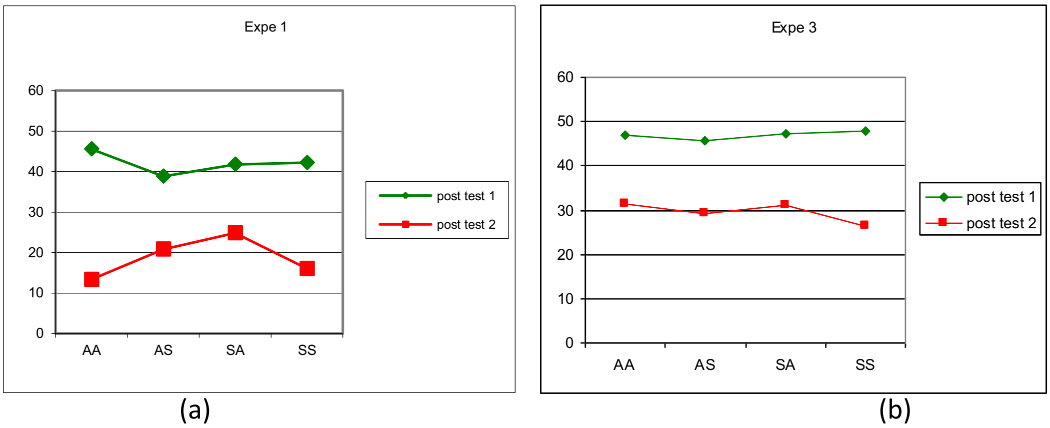


Figure 11.2. The result of comparison for post test 1 and post test 2 between (a) experiment 1 and (b) experiment 3.

Comparison of experiment 2 and 3

In order to compare experiment 2 and experiment 3, one ANCOVA analysis was performed with result for score the local movement (Post test 1), it showed no different between two of experiments, F (1,111) < 1, ns. Contrary in the functional mental model (Post test 2), participants in experiment 3 have lower score significantly than participants in experiment 2, F (1,111) = 15.16, p<0.001, η_p^2 = 0.12.

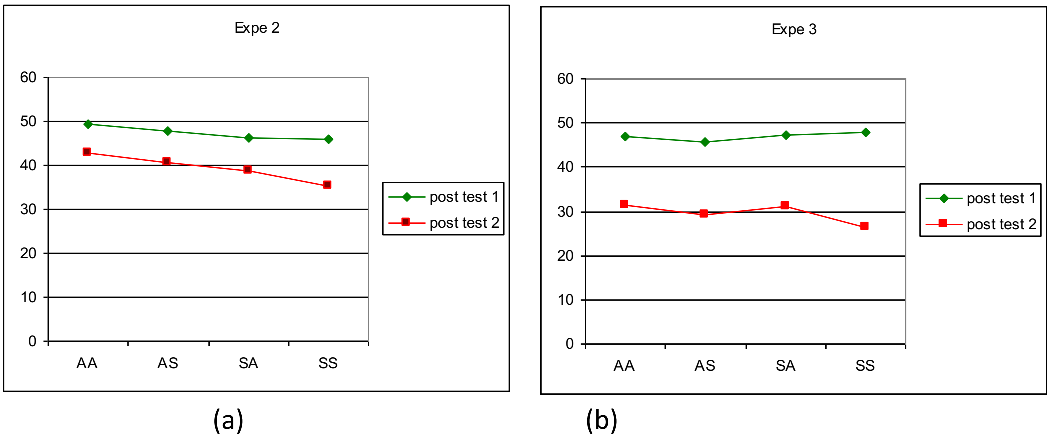


Figure 11.3. The result of comparison for post test 1 and post test 2 between (a) experiment 2 and (b) experiment 3.

To evaluate the effect of absence of the cueing in animation format representation in experiment 3, we compared the result for each tokens in all of experiment condition between experiment 2 and experiment 3 which is shown in Figure 11.4. A repeated measures ANOVA was conducted with four tokens scores as dependent variables and the experiment conditions and group level visual-spatial skills as categorical factor, showed a significant effect of token in each experiment conditions $F(3, 312) = 49.74$, $p < 0.001$, $\eta_p^2 = 0.32$.

In figure 11.4, the specific comparisons indicated that participants in experiment 3 obtained lower scores than participants in experiment 2, especially for the token 1 (key – whippen), in AS condition, $F(1,104) = 14.05$, $p<0.001$, and for token 3 (damper – hammer) in AA condition $F(1, 104) = 6.86$, $p<0.02$. There is no significant difference in SS and SA condition between two experiments.

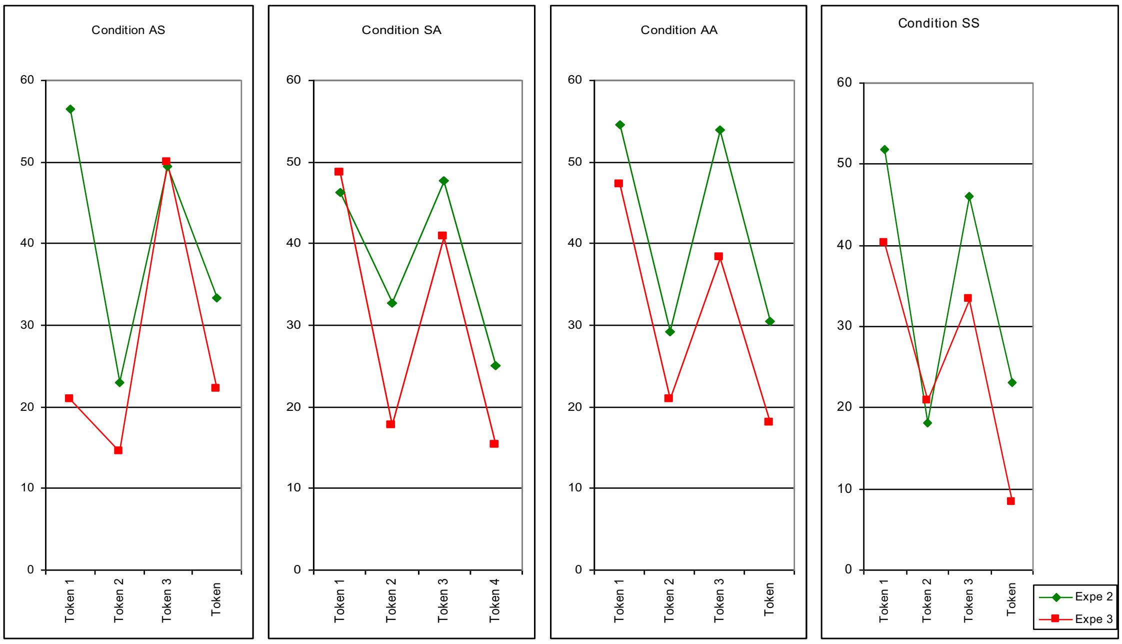


Figure 11.4. The result of comparison for each token in all of experiment condition between
(a) experiment 2 and (b) experiment 3
Note: token 1 (key – whippen), token 2 (hammer butt – jack), token 3 (damper – hammer), token 4 (back catch – back check).

Comparison result of eye movement

A comparison of the result of eye movement in experiment 2 and experiment 3, using an ANOVA repeated measures for each indicator showed no effects for the number of observations, $F(1,112) < 1$, ns, nor for the duration of fixation, $F(1,112) = 1.70$, $p=0.19$, and no effect for the number of fixation in each AOI, F (1,112) < 1, ns. However there is an interaction between factor experiment and the different AOI for the number of observation, F (3,336) = 5.16, p<0.002, η_p^2 = 0.04. Interaction was confirmed by the average duration of fixation in AOI-1 showed significantly higher in experiment 2 (17.78 ms *vs* 12.49 ms on average in experiment 3, p < 0.001). There is no significant difference for AOI-3 and AOI-4. For AOI-2 in contrast fixation times depend on a double interaction with the experimental conditions F (9,336) = 2.04, p<0.04, η_p^2 = 0.05.

11. 2 General Discussion

The present work was composed of four experiments focused on new segmentation methods, new cueing techniques and new format comparison. In experiment one, we make the hypothesis that dual format representation that is combining an integrated series of static pictures representation showing key steps of the dynamic process and a full animation representations was better than single format representation. We used series of static pictures format as simplified representations and animation format as dynamic representation allowing the learner to test the mental model which was previously built from the static pictures. Result showed that the dual format presentation resulted in higher comprehension than single format, especially for mental model building activities. Series of static pictures allowed the learner to apprehend the key stages in order to create a high quality mental model. Key frames provided learners with simplified representation more easily manipulate in working memory and facilitate the understanding of animation. When the animation was presented secondly, it strengthens the hypothetical mental model created previously.

For all experiment conditions, the local movement resulted in better scores than the functional mental model task. Especially in Animation and Animation (An+An) condition, the local movement was the highest because there are a lot of movement in the animation. For the functional mental model, the conditions of the dual format looked better than the condition with single format presentation. Single format presentation seemed to be superior for the local movement level. The series of static pictures served as “working model” to learn more from animation. According to the Animation Processing Model by Lowe and Boucheix (2008, 2011), there were four steps for learners when they process mental model building. The first was the selection process, the second was the decomposition stage which consisted of schematization with simplification and abstraction, the third stage composition of group of components by making inferences, with animation, and the fourth stage consisted in a validation stage.

In all our experiments the series of static pictures were used as a simplification and abstraction tool, the animation presentation was used as a tool composition for helping the inference ability between the static series of pictures. The learners had an opportunity to confirm their mental model constructed from the series of static pictures. Feedback from the animation would allow adjusting their representation previously built in the learning stage before.

In experiment two and three, we used the same dual format presentation in order to test new cueing techniques which consisted in signaling temporal tokens. Mayer & Chandler (2001) concluded that the 'segmenting' method decreases extraneous load for novices by breaking down the animations into segments corresponding to meaningful parts of demonstrated process and Betrancourt (2005) stated that the 'cueing' method decreases extraneous load by providing cues that focus the learners' attention on the relevant parts of the animations. The building of an efficient mental model might be easier when adding cues in color distributed of temporal tokens as a guide for attention direction on the causal chain of events.

First of all, as in experiment one, result for the local movement showed better score than result for the functional mental model. However, contrary with result in experiment one, for both the comprehension tests, there were no significant difference between the experimental conditions.
These compare with experiment one, in experiment two, the addition of spatial temporal cueing of tokens facilitated the recall of local movements and benefited also to the functional mental model building.

Based on the score for the functional mental model results indicated that participants were better in understanding on token 1 (key-whippen), token 3 (damper-hammer), token 2 (hammer butt-jack) and token 4 (balance -back check).

The understanding level for the three phases of the piano mechanism system was as follow. The best scores were found for phase 1 (strike phase): participants in all experiment conditions looked similarly to comprehend it, for phase 2 (rebound stage): participants who viewed animation followed by a series static pictures and animation-animation condition understood better than participants in other conditions, and in phase 3 (reset stage): participants had better scores when they started with static pictures.

Eye tracking results showed that the effect of the temporal spatial tokens was very clear on the direction of attention during the learning time. Analysis of the eye movement showed that participants took an advantage from dual format representation (SA and AS conditions). They showed better score from the number of observations than participants in the single format presentation (SS conditions).

From the result of eye tracking, it is clear that cued token directed efficiently the learners' attention when the cues appeared and until the cues disappeared. However, directing efficiently learners' attention did not result always in better comprehension. The comparison of experiment one and two, as well as the comparison between experiment one and three, showed that positive effect of cueing temporal spatial tokens not only on attention direction but also on comprehension.
However we found no significant difference between the cued group and the control group in experiment two. Further research need to be continued using temporal spatial tokens types.

Result of eye movement analysis on the animation following the series of static pictures revealed that when viewing animations with (Experiment two) or without cueing (Experiment three), participant preference was to look at the large elements with high diagram contrast of the piano mechanism.

However participants who viewed first the cued tokens of the series of static pictures before the presentation of the animation looked at more relevant component of the piano mechanism, which were signaled by the cued presented before. It is the case for example for the whippen, jack, and hammer butt. This result may be due to the lack of simplification of the information presentation during the first steps on learning process.

Another interesting point concerned the key salient balance-back check component. They attracted learners' attention soon and during a long time but this sustained attention did not result in better comprehension.

The executive functions test in experiment two and three were used to investigate the possible of effect of working memory capacity on understanding a complex mechanism system. We found that executive functions could play a role in animation processing. Result showed that The Stroop test as well as, N-Back test, and XO test could play a “moderator” effect between spatial ability and comprehension. However spatial ability effect on performance remains stronger than executive function test.

Another type of presentations 2D and 3D animation was used as another type of cueing. Indeed 3D presentation could change the perceptibility profile of the complex mechanical system. The eye tracking data indicated that the learners were focused more on relevant component of the piano mechanism (such as the whippen) in 3D presentation than in 2D.
Three D animation affected attention and comprehension of learners for depth. In the 3D version, whippen is more apprehendable than in 2D version.

This result is consistent with Boucheix and Lowe (2010) that approach of perceptibility profile of piano mechanism animation. Because of hammer is a large shaped item whose movement is rapid and covers much of the display area, so it has a high level of perceptibility. The results are similar too with those held by Pavclinic, Paul, Janet, and Tony (2001) who compared 2D and 3D animations in the understanding of stereochemistry processes. They supported that 3D animations with colourful molecular images encouraged students to practise and understand in Stereochemistry learning. Wu & Shah (2004) stated that presenting 3D animation was supposed to be a possible way of changing and improving students’ incomplete mental models.

Limitation of our research

This dissertation has taken a closer look about dual format presentation cueing as a way to foster learning from complex animation. In this research, all experiments were performed with the same system of the piano animation. This choice reinforced the consistency of our results over the four experiments of this dissertation. Also using the same system of the piano animation and the same method in our research, allowed us to compare between experiments in terms of alternate variables. However, our results could not be generalized to all types of an animation. Despite the limitation we can assume that the results could be generalized to other mechanical systems which have sequence of relevant events, organized long causal chains.

In our experiments, we used executive functions test to examine the possible effect of working memory on the understanding of a complex animation. This investigation was exploratory and executive functions were not in major focus of our research.

However, results from the test brought new insight on the news factors which could effect the cognitive processing of complex dynamic visualization. Further research in highly needed in this area, particularly about the relation between executive function test and visuo spatial ability tests.

Perspective for future research

It would be interesting to analyze how animations are processed in different content such as in chemical or biological processes. The dual format with animations should be effective for contents in which decomposition steps are difficult, and where the transitions between steps are important to understand. It could be interesting to extend the dual format presentation used in our experiment for the application in multidisciplinary field, by designing interactive features for learning scientific processes. Animation format would be a good visual display for people who interact with complex systems (e.g. in industrial and hospital).

We hope that our experiments could help designers to find out effective format presentation for learning from complex animation.

Recommendation for effective learning from a complex animation

This dissertation aimed to investigate the cognitive processing of animation and to find out which format of presentations could help the learner for the comprehension of complex systems. Nevertheless, the results could also be used as recommendations for the design of visual displays to facilitate the learners' of understanding.

According to our results, we can make the following recommendations:

(1) The use of dual format presentation can be as an effective way to comprehend dynamic process and help for developing dynamic mental model.

(2) The use of dual format with cueing could be an alternative which help learner in comprehension of learning complex system.

(3) The sequential format presentation allowed the learner to understand the key stages of dynamic process and create mental model that provide subject with simplified representation and facilitate the understanding of animation.

(4) 3D animation format could be an alternative which help the learner in order to foster learning by depth cues.

(5) Eye tracking measures could provide insight into how people understand and learn about complex systems and contribute to the understanding of the subject's perceptual processing during learning.

These recommendations should be carefully used, depending of the learning content which they can be applied. Maybe animation designing principle should be used in the way they are used in human machine area where design is driven by strict norms (ISO, Bastien, Scapin, and Leulier, 1999).

LIST OF REFERENCES

List of References

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

Agus Marco, Fabio Bettio, Enrico Gobbetti, Giovany Pintore. (2007). Medical visualization with New Generation Spatial 3D Displays, *Eurographis Italian Chapter Conference.*

Amadieu,F., Mariné,C., & Laimay,C. (2011). The attention-guiding effect and cognitive load in the comprehension of animations. *Computer in Human Behavior 27,* 36-40.

Ardac, D., & Akaygun, S. (2005). Using static and dynamic visuals to represent chemical change at molecular level. *International Journal of Science Education, 27(11),* 1269-1298. doi:10.1080/09500690500102284.

Arguel, A., & Jamet, E. (2009). Using video and static pictures to improve learning of procedural contents. *Computers in human behaviour, 25,* 354-359.

Ayres, P., & Paas, F. (2007). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology,21,* 695-700.

Baddeley, A.D. (1998). *Human memory : Theory and practice.* Boston: Allyn and Bacon.

Bastien, J.M.C., Scapin, D.L., & Leulier, C. (1999). The Ergonomic Criteria and the ISO 9241-10 Dialogue Principles: A pilot comparison in an evaluation task. *Interacting with computers, 11,* 299-322.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1973, 2002). *Differential aptitude test* (5^{th} ed.). Paris : ECPA.

Berney, S., & Bétrancourt.M. (2009). When and why does animation enhance learning? A meta-analysis. EARLI2009-Symposium “Designing effective dynamic and static visualizations in complex learning environment”.

Bétrancourt, M. and Tversky,B. (2000). Simple animations for organizing diagrams. *Human-Computer Studies.*

Bétrancourt, M. and Tversky,B. (2000). Effects of computer animation on users’performance: A review. *Le Travail Humain, 63,* 311-329.

Bétrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R.E.Mayer (Ed). *Cambridge handbook of multimedia learning.* 287-296. New York. Cambridge University Press.

Bétrancourt, M., Dillenbourg,P., Montarnal, C. (2003). *Computer technologies in powerful learning environments : the case of using animated and interactive graphics for teaching financial concepts*. In E. de Corte, L. Verschaffel, N. Entwistle, J. van

Merriënboer (Eds.), Unrevealing basic components and dimensions of powerful learning environments. 143-157.

Bétrancourt, M., Dillenbourg.P., Clavein, L. (2008). Display of Key Pictures from Animation : Effects on learning. *Understanding Multimedia Documents*. doi: 10.1007/978-0-387-73337. 61-78.

Boucheix, J.M. & Guignard, H. (2005). Which animation condition can improve test comprehension in children? *European Journal of Psychology of Education, 20(4),* 369-388.

Boucheix, J.M., & Lowe, R. K. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations.*Learning and instruction, 20*, 123-135.

Boucheix, J. M., & Schneider, E. (2009). Static and animated presentations in learning dynamic mechanical systems. *Learning and instruction, 19*, 621-642.

Canham, M., & Hegarty,M. (2010). Effects of knowledge and display design on comprehension of complex graphics? *Learning and Instruction*, *20,* 155- 166.

Carroll, J. (1993). Human cognitive abilities: A survey of factor analytical studies. New York: Cambridge University Press.

Casallas, J.S, (2009). Outils pour l'etude des processus cognitifs. *Mastere Specialise, Simulation,Realite Virtuelle,* Arts et Métiers Paristech.

Chi, M.T.H. (2000). Self-explaining expository texts: The dual process of generating inferences and repairing mental models. In R.Glaser (Ed), *Advances in Instructional Psychology: Volume: 5. Educational Design and Cognitive Science. 161-238.*

Chi, T.T.H., de Leeuw., Chiu, M.,& La Vancher,C. (1994). Eliciting self-explanations improves learning. *Cognitives Science, 18,* 439-478.

Cleveland, W.S. (1985). *The element of graphing data*. Monterey, CA: Wadsworth.

Cleveland, W.S., & Mc.Gill, R. (1984). Graphical perception: Theory, experimentation and application to the development of graphical methods. *Journal of the American Statistical Association, 79,* 531-554.

Cockburn, A., McKenzie, B. (2004). Evaluating spatial memory in two and three dimensions. *International Journal of Human-Computer Studies 61,* 359-373.

De Koning, B.B., Tabbers, H.K., Rikers, R.M.J.P., & Paas, F. (2007). Attention cueing as a means to enhance learning from an animation. *Applied Cognitive Psychology. 21:* 731-746.

De Koning, B.B., Tabbers, H.K., Rikers, R.M.J.P.,& Paas, F. (2009). Towards a framework for attentional cueing in instructional animations: Guidelines for research and design. *Educational Psychologycal Review 21: 113-140* DOI 10.1007/s10648-009-9098-7

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F (2010). Attention guidance in learning from a complex animation: seeing is understanding?. *Learning and instruction, 20*, 2, 111-122.

De Koning, B.B., Tabbers, H.K., Rikers, R.M.J.P., &Paas,F. (2011). Attention cueing in an instructional animation: The role of presentation speed. *Computer in Human Behavior 27,* 41-45

Denis, M., & de Vega, M. (1993). Modèles mentaux et imagerie mentale. In M. F. Ehrlich, H. Tardieu, & M. Cavazza. *Les modèles mentaux, approche cognitive des représentations*. Paris : Masson.

Duchowski, A.T. (2007). Eye tracking methodology: Theory and Practice. Second edition. Springer.

Faraday, P., & Sutcliffe,A. (1997). An empirical study of attending and comprehending multimedia presentations. *Proceedings of ACM Conference Multimedia 96 Conference,* 265-275.

Faraday, P., & Sutcliffe,A. (1997). Designing effective multimedia presentations. Proceedings of ACM conference on human factors in computing systems, 272-278.

Fischer,S., & Schwan, S. 2010.Comprehending animations: Effects of spatial cueing versus temporal scaling. *Learning and Instruction, 6,* 456-475.

Fischer,S., & Schwan, S. (2008). The role of speed variations in animations for attention and comprehension.

Fischer, S., Lowe, R., & Schwan, S. (2008). Effects of presentation speed of a dynamic visualization on the understanding of a mechanical system. A*pplied Cognitive Psychology, 22, 1126-1141.*

Freedmann, E.G.,& Shah, P. (2002). Toward a model of knowledge-based graph comprehension. In M.Hegarty, B.Mayer, & N.H. Narayanan (Eds.), *Diagrammatic representation and inference* (pp. 8-31). Berlin:Springer-Verlag.

Gavenda, J.D., Edington, J.R. (1997). Newton's Cradle and Scientific Explanation,*The Physics Teacher 35*. 411-417.

Hays, T.A. (1996). Spatial abilities and the effects of computer animation on short-term and long-term comprehension. *Journal of Educational Computing Research, 14,* 139-155.

Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. Journal of Experimental Psychology:Learning, Memory and Cognition, 18, 1084-1102.

Hegarty,M., Kriz.S & Cate. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction, 21(4),* 325-360.

Hegarty, M. (2004). Mechanical reasoning by mental animation. *Trends in Cognitive Science,* 8 (6), 280-285.

Hegarty, M. (2004). Dynamic visualizations and learning: Getting to the difficult questions. *Learning and Instruction, 15,* 343-351.

Hegarty, M & Waller, D.A. (2005). Individual differences in spatial abilities. *The Cambridge handbook of visuospatial thinking*. Cambridge University Press. 121-169.

Hegarty, M. (2005). Multimedia learning about physical systems. In R.E. Mayer (Ed), *The Cambridge handbook of multimedia learning.* 447-465. Cambridge, England: Cambridge University Press.

Hegarty, M. (2011). The cognitive science of visual-spatial displays: Implications for design. *Topics in Cognitive Science 3*. 446-474. doi: 10.1111/j.1756-8765.2011.01150.x

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

Hegarty, M., & Kriz, S. (2008). Effects of knowledge and spatial ability on learning from animation. In R. Lowe, & W. Schnotz (Eds.). *Learning with animation*. New York : Cambridge University Press.

Höffler, T.N., & Leutner,D., (2007). Instructional animation versus static pictures: a meta analysis. *Learning and Instruction, 17,*722-738

Höffler, T. N. (2010). Spatial ability: its influence on learning with visualizations – a meta-analytic review. *Educational psychological review, 22*: 245-269.

Höffler, T. N., & Leutner, D. (2010). The role of spatial ability in learning from instructional animations–Evidence for an ability-as-compensator hypothesis. *Computers in human behaviour*, in press.

Hollands, Justin G and Ivanovic Nada. (2002). Task switching with 2D and 3D displays of geographic terrain: The role of visual momentum. Paper presented at the RTO IST

Workshop on "Massive military data fusion and visualisation: Users talk with developers", Norway, 10-13 September 2002.

Huk, T., Steinke, M.,& Floto, C. (2003). *Helping teachers developing computer animations for improving learning in science education.* Albuquerque, USA: Proceedings of the Society for Information Technology and Teacher Education.

Huk,T. (2006). Who benefits from learning with 3D models?The case of spatial abilty. *Journal of Computer Assisted Learning 22*, 392-404. Göttingen, Germany.

Huk,T., Steinke,M., Floto,C. (2010).The educational value of visual cues and 3D-representaional format in a computer animation under restricted and realistic conditions. *Instr Sci*. DOI 10.1007/s11251-009-9116-7. Springer Science+Businnes Media B.V

Hyönä, J. (2010).The use of eye movements in the study of multimedia learning. *Learning and Instruction*, *20,* 172-176.

Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction this issue*

Johnson-Laird, P.N. (1983). Mental Models : Toward a cognitive science of language, inference and consciousness. Cambridge University Press.

Johnson-Laird, P. N. (1993). La théorie des modèles mentaux. In M. F. Ehrlich, H. Tardieu, & M. Cavazza. *Les modèles mentaux, approche cognitive des représentations*. Paris: Masson.

Johnson-Laird, P. N.(1998).Imagery, visualization, and thinking.In J. Hochberg (Ed.), Perception and Cognition at the Century's End. 441-467. San Diego, CA: Academic Press

Jamet, E., Gavota, M., and Quaireau, C. (2008).Attention guiding in multimedia learning. *Learning and Instruction. 18,* 135-145.

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

Kosslyn, S.M. (2006). *Graph design for the eye and mind.* New york: Oxford University Press.

Kombartzk, U., Ploetzner, R., Schlag, S., & Metz, B. (2010). Developing and evaluating a strategy for learning from animations. *Learning and Instruction, 20,* 424-433.

Kriz, S., & Hegarty, M.(2007). Top-down and bottom-up influences on learning from animation. *International Journal of Human Computer Studies, 65*, 911-930

Large, J.A., Beheshti, J., Breuleux, A., & Renaud, A. (1998). The effect of animation in enhancing descriptive and prosedural texts in a multimedia learning environment. *Journal of American Society for Information Science, 47,* 437-448.

Lightner, N. (2001). Model testing of users' comprehension in graphical animation: The effects of speed and focus areas. *International Journal of Human-Computer Interaction, 13(1),* 53-73.

Loman, N.L.,& Mayer, R.E. (1983). Signaling techniques that increases the understandability of expository prose. *Journal of Educational Psychology, 75,* 402-412.

Lorch, R.F.(1989). Text signaling devices and their effects on reading and memory processes. *Educational Psychological Review, 1,* 209-234.

Lorch, R.F., & Lorch, E.P. (1996). Effects of organizational signals on free recall off expository text. *Journal of Educational Psychology, 88,* 38-48.

Lowe, R.K., (1999).Extracting information from an animation during complex visual learning. *European Journal of Psychology of education, 14,* 225-244.

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

Lowe, R.K., (2004).Interrogation of a dynamic visualization during learning. *Learning and Instruction*, *14,* 257- 274

Lowe, R. K. & Boucheix, J.M. (2007). Eye tracking as a basis for improving animation design. Contribution to the symposium: Eye tracking as a means for detailed analyses of multimedia learning processes. *European Association for Research on Learning and Instruction conference.* Budapest.

Lowe, R. K., & Boucheix, J.M. (2008). Learning from animated diagrams: How are mental models built ?. In G. Stapelton, J. Howse, & J. Lee (Eds.). *Theory and applications of diagrams*. Berlin : Springer.

Lowe, R.K., & Boucheix, J.M. (2008). Supporting relational processing in complex animated diagram. In G. Stapelton, J. Howse, & J. Lee (Eds.): *Diagram 2008*, 391-394. Springer-Verlag, Berlin Heidelberg.

Lowe, R.K., & Boucheix, J.M. (2011). Cueing complex animations: Does direction of attention foster learning processes*? Learning and Instruction, 21,* 650-663.

Mautone, P.D. & Mayer, R.E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93,* 377-389.

Mayer. R.E. (1989). Models for understanding. *Review of Educational Research, 59 (1),* 43-64.

Mayer, R.E. (2001). Multimedia Learning. Cambridge University Press.

Mayer, R.E. (2005). *The Handbook of Multimedia Learning*, New York: Cambridge University Press.

Mayer, R. E. (2008). Research-based principles for learning with animation. In R. Lowe, & W. Schnotz (Eds.). *Learning with animation*. New York : Cambridge University Press.

Mayer,R.E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction*, *20, 167-171.*

Mayer, R.E. and Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93(2),* 390-397

Mayer, R.E., Hegarty.,& Campbell, J., (2005).When static media promote active learning: annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied, 11,* 256-265.

Mayer, R.E., Heiser, H., & Lonn, S. (2001). Cognitive constrains on multimedia learning: when presenting more material results in less understanding. *Journal of Educational Psychology, 93,* 187-198.

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

Mayer, R.E and Moreno,R. (2002). Animation as an Aid to Multimedia Learning**.** *Educational Psychology Review, Vol.14, No 1.*

Meyer, K., Rasch, T., & Schnotz,W. (2010). Effects of animation's speed of presentation on perceptual processing and learning. *Learning and Instruction, 20*, 136-145.

Mc Clean, P., Johnson, C., Rogers, R., Daniels, L., Reber, J., Slator, B.M. (2005). Molecular and cellular biology animations: Development and impact on student learning. *Cell Biology Education*, 4, 169-179 doi:10.1187/cbe.04-07-0047

Miyake, N. (1986).Constructive interaction and the iterative process of understanding. *Cognitive science*, *10*, 151-177.

Morand,L.,& Bétrancourt, M. (2010). Collaborative learning with single or multiple animations. *EARLI Conference, SIG2*. *Comprehension of text and graphics, Tübingen*, August 25-28.

Moreno, R. (2007). Optimising learning from animation by minimising cognitive load: Cognitive and affective consequences of signalling and segmentation methods. *Applied Cognitive Psychology. 21*: 765-781.

Narayanan, N.H.,& Hegarty, M. (1998). On designing comprehensible interactive hypermedia manuals. *International Journal of Human Computer Studies, 48,* 267-301.

Narayanan, N.H.,& Hegarty, M. (2002).Multimedia design for communication of dynamic information. *International Journal of Human Computer Studies, 57,* 279-315.

Ozcelik,E., Ari,A.I., Cagiltay,K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior 26,* 110-117.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38,* 1-4.

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

Park, O.C., & Hopkins, R., (1993). Instructional conditions for using dynamic visual displays: a review. *Instructional Science, 21*, 427-449.

Pavclinic, S., Buckley, P., Davies, J., and Wright, T. (2001).Computing in stereochemistry-2D or 3D representations? *Research in Science Education-Past, Present, and Future, 295-300.* Kluwer Academic Publishers. Printed in the Netherlands.

Pinker, S. (1990) graph comprehension. In R.Freedle (Ed). Artificial Intelligence and the future of testing. 73-126.

Poole, Alex and Ball, Linden J. (2005).Eye tracking in human-computer interaction and usability research: Current Status and Future Prospects.

Putri, D.K., Boucheix, J.M, & Casallas, J.S. (2009). “2D and 3D animation in learning complex mechanism system: an eye tracking investigation” Forum des Jeunes Chercheurs, Dijon, Université de Bourgogne.

Rayner.,K. (1998). Eye movements in reading and information processing: 20 years of *research. Psychological Bulletin*, 124(3), 372-422.

Rebetez, C., Bétrancourt, M., Sangin, M., & Dillenbourg, P. (2010). Learning from animation enabled by collaboration. *Instructional science, 38*, 5, 471-485.

Rieber,L.P. (1990). Using computer animated graphics with science instruction with children. *Journal of Educational Psychology*, *82,* 135-140.

Salvucci, Dario D. and Goldberg, Joseph H. (2000). Identifying Fixations and Saccades in Eye-Tracking Protocols. Palm Beach Gardens, Florida, United States. *Proceedings of the symposium on eye tracking research & applications*, 71-78.

Schwan, S. & Riempp, R. (2004). The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction, 14,* 293-305.

Schnotz, W. (2001). Sign system, 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. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*. 14(1), 101-120.

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

Schnotz,W., & Rasch,T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: why reduction of cognitive load can have negative results on learning. Educational Technology Research and Development, 53, 47-58.

Schnotz,W., & Rasch. (2008).Functions of animation in comprehension and learning. In R.K.Lowe, & W. Schnotz (Eds.), Learning with animation: Research implications for design pp. 92-113., New York, Cambridge University Press.

Schnotz, W., & Lowe, R. (2008).A unified view of learning from animated and static graphics. In R. Lowe, & W. Schnotz (Eds.). *Learning with animation*. New York : Cambridge University Press.

Sims, V.K., & Hegarty, M. (2007).Mental animation in the visuospatial sketchpad: Evidence from dual-task studies. *Memory & Cognition, 25*, 321-332.

Spanjers, I. A. E., Wouters, P., van Gog, T. & van Merriënboer J. J. G. (2010).An expertise reversal effect of segmentation in learning from animated worked-out examples. C*omputers in human behaviour*, in press.

Spector, A., & Biederman, I.(1976). Mental set and mental shift revisited. *American journal of psychology, 89*, 4, 669-679.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions.*Journal of experimental psychology, 18*, 643-662.

Smallman, H., John, M., Oonk.H., and Cowen. (2000).When beauty is only skin deep: 3-D realistic icons are harder to identify than conventional 2-D military symbols. *Proceedings of the IEA 2000/HFES 2000 Congress*

Smallman, H., John, M., Oonk, H., and Cowen. (2001) Information Availability in 2D and 3D displays. *IEEE Computer Graphics and Applications, 21*, 51-57.

Smallman, H., Cowen, M.B, and Oonk, H. (2001). The use of 2D and 3D displays for shape understanding versus relative-position tasks. *Human Factors, 43,* 79-98.

Sweller, J. (1999). Instructional design in technical areas. Camberwell. Australia: ACER Press.

Sweller, J.(2005).Implications of cognitive load theory for multimedia learning. In R.E.Mayer (Ed).*Cambridge handbook of multimedia learning*. 287-296. New York. Cambridge University Press.

Sweller,J.,& Chandler,P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12,* 185-223.

Robertson, G., Czerwinski, M., Fisher, D., & Lee, B. (2009). Selected human factors issues in information visualization. *Reviews of Human Factors and Engineering, 5*, 41-81.

Tabbers, H., Martens, R.,& Van Merriënboer, J. (2004). Multimedia instructions and Cognitive Load Theory: Split-attention and modality effects. Paper presented at the AECT 2000 in Long Beach. California.

Tversky, B., Morrison, J.B., & Bétrancourt, M., (2002). Animation: can it facilitate? *International Journal of Human Computer Studies,57,247-262*

Tversky, B., Heiser, J., Mackenzie, R., Lozano, S., & Morrison, J. (2008). Enriching animations. In R. Lowe, & W. Schnotz (Eds.). *Learning with animation*. New York : Cambridge University Press.

Van Oostendorp, H., & Beijersbergen, M.J. (2007). *Animated Diagram: Their effect on understanding confidence and mental effort.* Paper presented at the Bi-annual meeting of the European Association of Research on Learning and Instruction. Budapest, Hungaria.

Wiebe. E., Slykhuis.D., Madison. J., and Savage, D. (2004). Answering question with 2D and 3D topographic maps: use of eye tracking to understand and representational reasoning, *Proceeding of the NARST* (2005) Annual Meeting, Dallas United States.

Weigand, F.S., Kohnert,A., & Glowalla,U. (2009). A closer look at split visual attention in system and self-paced instruction in multimedia learning. *Learning and Instruction, 1-11.Article in Press.*

Weigand, F.S., & Scheiter, K. (2010). The role of spatial descriptions in learning from multimedia. *Computer in Human Behaviour. Article in Press*

Wu, H.K., Krajcik, J.S., & Soloway, E. (2001). Promoting understanding of chemical representations: Students' use a visualization tool in the classroom. *Journal of Research in Science Teaching, 38(7),* 821-842.

Yarden. H and Yarden. A, (2010). Leaning using dynamic and static visualizations: *Students' comprehension, prior knowledge and conceptual status of a biotechnological method.* 40:375-402 doi: 10.1007/s11 165-009-9126-0. Springer Science., Business Media.

Zacks, J.M, & Tversky,B.(2001).Event structure in perception and conception. *Psychological Bulletin, 127,* 3-21.

Zack, J.M. (2004). Using movement and intensions to understand simple events.*Cognitive Science 28*, 979-1008.

Zacks, J.M., Speer, N.K.,Swallow, K.M., Braver, T.S., and Reynolds, J.R. (2007). Event perception: A mind brain perspective. *Psychological Bulletin, Vol 133, No.2*, 273-293. *The American Psychological Association*. doi: 10.1037/0033-2909.133.2.273