A psychology literature study on modality related issues for multimodal presentation in crisis management

Yujia Cao
Human Media Interaction (HMI)
University of Twente
7500 AE Enschede, the Netherlands
y.cao@utwente.nl

Abstract

The motivation of this psychology literature study is to obtain modality related guidelines for real-time information presentation in crisis management environment. The crisis management task is usually companied by time urgency, risk, uncertainty, and high information density. Decision makers (crisis managers) might undergo cognitive overload and tend to show biases in their performances. Therefore, the on-going crisis event needs to be presented in a manner that enhances perception, assists diagnosis, and prevents cognitive overload. To this end, this study looked into the modality effects on perception, cognitive load, working memory, learning, and attention. Selected topics include working memory, dual-coding theory, cognitive load theory, multimedia learning, and attention. The findings are several modality usage guidelines which may lead to more efficient use of the user's cognitive capacity and enhance the information perception.

Content

1 Introduction

This psychology literature study is part of a PhD research on **multimodal presentation for preventing tunnel vision in crisis management**. In this section, a brief introduction on the research topic and the current progress will be given (see [1] for a more elaborated description and related references). The motivation and scope of this literature review will also be discussed at the end of this section.

Crisis management is a strategic management activity which aims to prevent or minimize the negative impact of a crisis. It is a challenging and stressful task because time is critical and people involved typically have to deal with high risk, high information density, and high uncertainty. This is especially true for the decision making group (the crisis managers) whose decisions and commands are crucial. Under stress and high working load, the decision makers often have the tendency to unintentionally bias their information perception and problem solving processes. We have defined this phenomenon as tunnel vision. Tunnel vision includes two aspects:

1) framing bias

Crisis managers tend to create one coherent description of the crisis event without full awareness of what has actually happened. They also tend to rely on standard operating procedures, previous ways of understanding without reexamination.

2) confirmation bias

Crisis managers tend to perceive and confirm only clear and familiar information which aligns with their subjective understanding and explanation of the crisis situation. Once a solution is forming, they tend to stick to it without examining other possibilities.

The two aspects of tunnel vision may lead to growing biases in the diagnosis of the real crisis situation, which in turn cause costly delay and loss. Aiming at preventing tunnel vision, an information assistant system has been proposed by the author [1]. Briefly speaking, this system serves as a platform for monitoring the on-going crisis event. It captures all real-world communications and multimodal signals, applies fusion, reasoning and filtering technologies, and eventually presents them to the users (crisis management located in the crisis response center). During these processes, the information quality is enhanced and the quantity is reduced. The research focus of the author is the design of the multimodal presentation module which aims at forming an effective and efficient presentation of the on-going crisis event. An effective and efficient presentation manner is supposed to prevent the user's cognitive overload and enhance the user's information perception performance. It is expected that, when working with such an information assistant system, the user is more likely to stay aware of the actual situation, stay open minded to possibilities, and make proper decisions.

As mentioned above, our goal is to achieve an effective and efficient presentation manner which prevents cognitive overload and assists information perception of the user. One of the factors which we believe to have major influences on the presentation performance is the modality usage. Accordingly, the research question is: how should available modalities be allocated and combined for presenting each piece of information, in order to make efficient use of the user's cognitive capacity and make sure that the information is actually perceived by the user. Modality planning approaches in the AI domain mostly intend to achieve a better representational performance, i.e. to better convey the meaning of the information. However, due to the specific application, the perceptional properties of modalities should also be taken as a very important concern. Therefore, a literature study into the psychology domain has been carried out, in order to obtain knowledge about the modality effect on cognitive load and attention.

2 Overview of this literature study

There is a huge body of psychology studies which addresses the effect of modality on perception, cognitive load, working memory, learning, and attention. They spread into many branches, including educational psychology, cognitive psychology, experimental psychology and neuropsychology. In this literature review, we intend to find principles for two aspects in modality planning:

1) Modality effects on cognitive load and the occupation of working memory

In the crisis situation, due to the time urgency and information overload, crisis managers may undergo cognitive overload. However, the emergency response tasks require available cognitive capacity for situation diagnose and decision making. Therefore, it is desired that our information presentation manner helps to reduce the cognitive load and make more efficient use of the working memory. From psychology literatures on cognitive load theory, working memory theory and dual coding theory, we expect to have better understanding on human being's cognition system and the working memory structure, and to obtain the modality effects on cognitive load and the occupation of working memory. Eventually, we intend to extract modality usage principles which lead to a better use of the user's cognitive capacity. Guidelines on the opposite side, i.e. the modality usages which may bring high cognitive load, are also valuable for this research.

2) Modality effects on attention

From the studies on attention, we expect to obtain modality-related findings which provide answers for the question: Can we adjust the attention-capture ability of the presentation by the usage of modalities? The question arises from the concern that the user may fail to perceive the newly-presented information, especially when he is concentrating on a certain previously-presented issue. When the user is not experiencing cognitive overload, we expect to be able to direct the user's attention to each piece of newly-presented information. Under the assumption that the user is attending the latest presented issues, the modality planning for the next piece of information should take the currently used modalities into account, especially when the location of the new information on the display is likely to be outside the user's attended area. When cognitive overload occurs, the modality usage of all information on the display should make contrast between the most urgent/important issues and the rest. In summary, we expect to gather knowledge about how attention is related to the modality usage.

According to our application, we will focus on visual, auditory, verbal, and non-verbal properties of modalities. Less relevant properties, such as tactile, olfactory, and gustatory, are out of the scope of this study. The selected topics include 1) working memory theory, 2) dual-coding theory (mental representation), 3) cognitive load theory, 4) multimedia learning theory, and 5) attention. The following sections give the review of selected literatures and summarize the findings.

3 Working memory and dual-coding theory

Working memory (or short-term memory) is a theoretical construct in cognitive psychology that refers to the system involved in the temporary maintenance and manipulation of information [2]. Long-term memory is a system that allows long-term (maybe a life-time) information store and retrieval. The limitation of working memory capacity has been well known since Miller [3]. Only a limited amount of elements can be stored and processed in working memory at the same time. In contrast, an unlimited number of elements can be held in long-term memory in the form of hierarchically organized schemas. During a high work load task, such as crisis management, working memory is more likely to be a cognitive bottleneck that limits the performance. Therefore, it is necessary to look at the current understandings on how working memory works.

Baddeley and Hitch [4] have proposed a very famous multi-component working memory model, see figure 1. This model contains a central executive system aided by two subsidiary systems – a visual-spatial sketch pad and a phonological loop. The phonological loop has a phonological store for temporarily storing auditory information. It also includes a rehearsal system. Auditory traces within the store were assumed to decay over a period of about two second unless being refreshed by the rehearsal system. The visuospatial sketchpad is assumed to temporarily maintain visual information and to form a relation between visual and spatial information. The information stored in the two subsidiary systems is retrieved by the central executive system, which is assumed to be an attentional system whose role extends beyond memory functions. The central executive is believed to involve in (at least) attention management, new learning, comprehension, and reasoning. As the name indicated, it might serve as a center processing and control system. This model has been supported by psychological, physiological, and anatomical evidences [5]. Each component in this model is assumed to have limited capacity, therefore the total capacity of working memory is also limited [6].

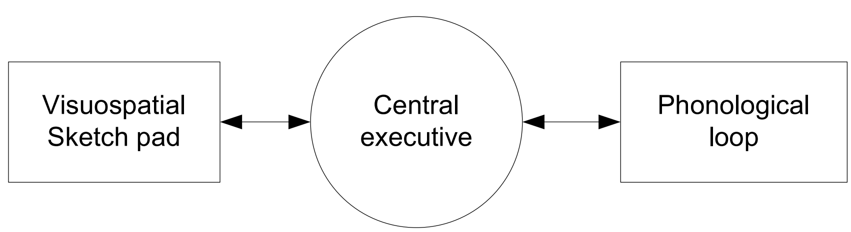

Figure 1: a multi-component working memory model proposed by Baddeley

Several conclusions can be extracted from the working memory theory for this research:

1) The working memory has separated stores for visual information and auditory information. Therefore, in order to make better use of the working memory

capacity, it is preferable to let the two perception channels share the presentation task in a coordinated manner.

2) Each perception channel (visual and auditory) has limited capacity. The total amount of information that needs to be perceived in each channel shouldn't exceed the corresponding capacity. Attention-related research suggested that the visual memory is limited typically to about three or four objects [7, 8] (as cited in [5]). This capacity limitation is one of the causes of the change blindness phenomenon, whereby objects in scenes can change color, move or disappear without people noticing (see section 5 for more detail).

The central executive is the most important but least understood component of working memory. Little is known about how information perceived and stored in the two subsidiary systems is processed by the central executive to serve different tasks. However, the dual coding theory from Paivio [9] provides more knowledge on human being's mental structures and processes. According to the dual coding theory, mental processes are dynamic associative processes that operate on a rich network of modality-specific verbal and nonverbal representations (see figure 2). Verbal representations contain visual, auditory, and other modality-specific verbal codes. Nonverbal representations include images, environmental sounds, actions, and other nonlinguistic objects and events. Two mental systems process verbal and nonverbal representations respectively. The associative connections are the links that join representations within each system. The two systems are further linked into a more complex dynamic network through referential connections. The referential connections join corresponding verbal and nonverbal codes into knowledge that can be acted upon, stored, and retrieved for subsequent use.

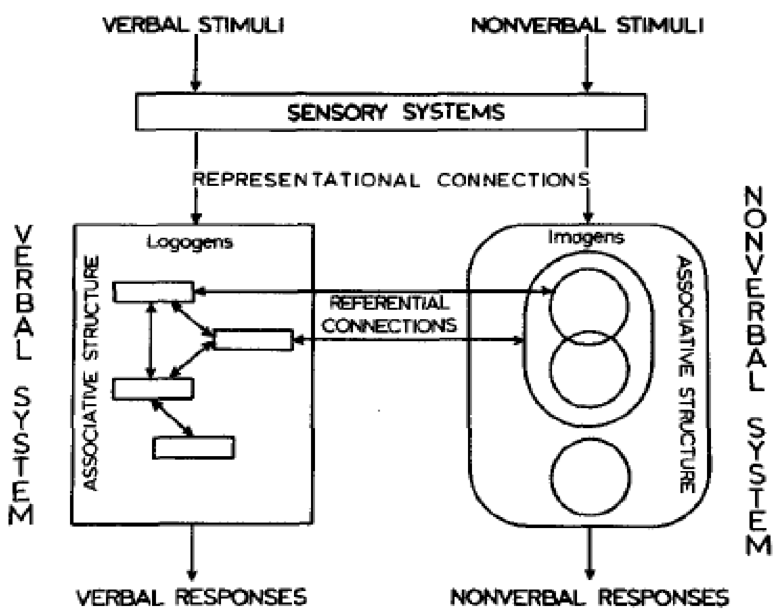

Figure 2: Verbal and nonverbal symbolic systems of Dual Coding Theory (from [9])

The dual coding theory is empirically well-founded. It has been applied to explain specific phenomena in many areas of education and at many levels [10]. It is demonstrated that verbal and nonverbal associative processes play major roles in various educational domains, such as knowledge comprehension, memorization, effective instruction design, the learning of motor skills, etc. Educational materials which contain associated verbal and nonverbal codes and effectively activate verbal/nonverbal associative/referential processes normally lead to better performance in comprehension, learning and memorization. Although the task of presenting a crisis scenario is different from giving educational instruction, it is still reasonable to assume that presenting information with both verbal and non-verbal modalities in a well-associated manner may lead to better situation memorization and diagnose. A well-associated manner means that it doesn't require much cognitive effort to build referential connections between verbal and nonverbal presentations.

Although these aforementioned two theories haven't been explicitly related to each other, it seems reasonable to interpret that the dual-coding structure is embedded in the central executive. When the stored visual information and auditory information are retrieved into the central executive, they are represented and re-classified by another feature (verbal or nonverbal). The central executive is also responsible for building associative and referential connections, although this is by no means its only function. As working memory has limited capacity, the capacity of each system is limited. The cognitive resource for making referential and associative connections is also limited.

4 Cognitive load theory and multimedia learning

Cognitive Load is a term that refers to the load on working memory during problem solving, thinking and reasoning [11]. Cognitive load theory has been constructed to describe the structure of human cognition, serving as a theoretical foundation for the design of educational instructions [11-13]. Based on the assumption that the amount of cognitive capacity available is limited (working memory capacity is limited), the cognitive load theory suggests that the total amount of cognitive load for a particular learner under particular condition is defined as the sum of three types of load: one type that is attributed to the inherent structure and complexity of the instructional materials and cannot be influenced by the instructional designer (intrinsic load), and two types that are imposed by the requirements of the instruction and can therefore, be manipulated by the instructional designer (extraneous and germane load). The intrinsic load is determined by an interaction between the nature of the material being learned and the expertise of the learners, therefore it cannot be directly influenced by instructional designers. The extraneous load is the extra load beyond the intrinsic load resulting from mainly poorly designed instruction. The germane load is induced by learner's mental effort to process and comprehend the material. The difference between the total cognitive load and the processing capacity of working memory is defined as free cognitive resource [14]. High cognitive load occurs when free cognitive resource approaches zero. Based on the cognitive load theory, the goal of instructional design is to adjust the extraneous load and the germane load the instructional materials may impose upon the learners and reduce the cognitive load required for learning.

Many research efforts have been investigated in the design of multimedia learning materials which bring cognitive load benefits for learners [15-28]. The relations between the usages of modalities (the modality effect) and the demand of cognitive load have been a particular interest in this research area. Although the goal of this research differs from improving learning, the assumptions and principles in the multimedia learning research are still highly relevant in the sense that both researches aim at designing multimodal presentations which reduce the cognitive load of the users. The cognitive theory of multimedia learning is based on four fundamental assumptions [14, 22].

1) The dual-channel assumption

This assumption refers to the sensory modality of information perception that visual information and auditory information are perceived in different channels that correspond to the visuospatial sketchpad and the phonological loops from Baddeley's working memory model [4].

2) The dual-coding assumption

This assumption is taken from Paivio's dual-coding theory [9]. It refers to the mental representation modes and indicates that verbal and nonverbal materials are processed and mentally represented in separate but interconnected systems.

3) The limited capacity assumption

This assumption posits that the amount of processing that can take place within each information perception channel and processing system is extremely limited, taken from the Sweller's cognitive load theory [12] and the Baddeley's working memory theory [4].

4) The active learning assumption

This assumption suggests that meaningful learning occurs when learners engage in active cognitive processing including paying attention to relevant incoming words and pictures, mentally organizing them into coherent verbal and nonverbal representations, and mentally integrating verbal and nonverbal representations with each other and with prior knowledge.

The research findings in the multimedia learning domain by Mayer et al. can be summarized as 9 effects [23, 26], including the modality effect, the contiguity effect, the multimedia effect, the personalization effect, the coherence effect, the redundancy effect, the pre-training effect, the signaling effect, and the pacing effect. Due to different research backgrounds, not all effects are relevant to this research. Only the relevant ones are introduced here.

1) The multimedia effect

The multimedia effect [20-23, 26, 29] shows that students learn better from a multimedia explanation presented in coordinated verbal (written text or narration) and nonverbal (pictures or animations) formats then in verbal format alone. This effect is consistent with the dual-coding theory. Associated verbal and nonverbal materials can activate the student's mental activities to build referential connections and associative connections; therefore they learn more deeply. The multimedia effect is stronger for low prior knowledge learners than for high prior knowledge learners [20]. This is because the high prior knowledge learners are more likely to be able to independently create useful mental images solely from the verbal materials. Therefore they benefit less from a coordinated verbal and nonverbal presentation.

2) The modality effect

When a learning material contains animation and text which are presented in different areas on the screen, students have to split his visual attention to animation and text. When they are reading the text, they may miss some part of the animation, and vice versa. This effect is named as the split-attention effect [18, 19, 25]. The modality effect [18, 19, 21, 23, 25, 26, 29] points out a solution to this problem which is to present the verbal information (text) with an auditory modality rather than a visual modality. A series of experiments proved that students perform better on tests of problem-solving when scientific explanations were presented as animation and narration, rather than as animation and on-screen text. This result can be explained by the dual-channel assumption. Animation and on-screen text together may impose high load on the visual channel. By replacing text by narration, part of the perception load is carried by the auditory channel. More effective capacity of the working memory is used. Therefore the learners perform better at information perception and further processing.

Two other methods have been proposed to reduce the split-attention effect, in the situation that the verbal and nonverbal information both have to be presented in the visual form. One is called color coding method [18] which is to color the related verbal and nonverbal elements with the same unique color. The other one is to physically locate the related elements near to each other (see the contiguity effect). Both of these two alternative techniques intend to reduce the load on the visual channel and assist the building of the referential connections, and eventually circumvent the cognitive load consequences of the split-attention effect.

3) The contiguity effect

The contiguity effect [20, 21, 23, 26-29] consists of two aspects - a spatial-contiguity effect and a temporal-contiguity effect. The spatial-contiguity effect refers to learning enhancement when corresponding verbal and nonverbal information (both visual) are physically well-integrated rather than physically separated. This is another solution for reducing the split-attention effect. The physical integration reduces the part of load which is needed for visual searches and mappings in the visual channel. Therefore the visual channel overload is less likely to happen and the students can perceive the information more effectively. The physical integration also assists the referential connections between corresponding verbal and nonverbal information which need to be integrated, and results in more free cognitive resources for comprehension and problem solving. The temporal-contiguity effect refers to learning enhancement when corresponding visual and auditory materials are temporally synchronized, i.e. presented simultaneously rather than successively. The explanation is also that the temporal synchronization assists the referential connections between verbal and nonverbal information. The learner doesn't have to hold one part of the information in the working memory and wait for other corresponding contents which need to be integrated.

In general, the integrated and synchronized corresponding verbal and nonverbal information facilitate the reference connections and enhance learning. The application of the contiguity effect has some conditions. First, it has been proved that students with high spatial ability can benefit more from the contiguity effect, because building the associative connections among visual representations is relatively effortless for them, therefore they have more free cognitive resources for building the referential connections [27]. Second, the contiguity effect doesn't hold for the information sources which don't have to be integrated in order to be understood. A long-term study has shown that integrated instructions are not effective in areas where verbal or nonverbal information sources are completely intelligible in isolation [16].

4) The coherent effect

The coherent principle [21-24, 26, 28] says that students learn better from a multimedia explanation without extraneous materials rather than with extraneous materials. Many experiments have provided consistent evidences for this principle, and it is also very well understandable from common sense. Irrelevant materials take up part of the limited cognitive capacity and the students need to devote extra cognitive resource to distinguish the irrelevant materials.

5) The redundancy effect

The redundancy effect refers to the fact that students understand a multimedia presentation better when words are presented as narration rather than as narration and on-screen text. [18, 21, 23, 24, 26, 28, 29]. This effect indicates that presenting the same information content in many formats, although may guarantee the perception, does not always enhance comprehension and learning. The benefit of multimodality presentation has some conditions. More modalities are not always better than less. The redundancy information may cause the overload of the perception channels and require extra mental resource for building referential connections.

6) The personalization effect

The personalization principle [21-23, 26] indicates that students learn better when the words are presented in a conversational style rather than a formal style. The explanation might be that a conversational style may make the learners more willing to accept what they are being told. Therefore, students work harder to understand an explanation when they are personally involved in a conversation.

All the aforementioned effects are supposed to be stronger when the contents of the presented materials are in nature more difficult to comprehend and/or the problem-solving tasks during learning are more complex. Difficult materials result in higher intrinsic cognitive load. Complex tasks require more free cognitive resources for reasoning and problem-solving. In these high-cognitive-load-requirement situations, the free cognitive capacity is more crucial to performance and cognitive overload is more likely to occur. In contrast, if the material comprehension and problem-solving are both relatively easy and don't require much cognitive effort in total, the aforementioned effects might be very weak or even not visible because the working memory capacity can afford the unnecessary working load.

Finally, the long term memory can also influence the burden on working memory [18]. Relevant knowledge stored in the long-term memory can be retrieved to facilitating the associative and referential connections among new information. Therefore, the cognitive load during learning and problem-solving can also be reduced by trainings and experiences.

5 Attention

The attention-related literature study started from looking at the change blindness phenomenon, which refers to the striking failure to observe large changes in a scene (see section 5.1). The change blindness phenomenon is worth our attention, because it would be a serious failure if some of the presented information on the display goes unnoticed. Driven by this concern, we turned to the research topic of attention guidance. The intention is to figure out what are the factors which guide attention. There is a huge body of attention related research, from single-modality attention (e.g. visual attention) to cross-modality attention. According to our application, our literature selection focuses on visual attention guidance (section 5.2) and cross-modality attention (section 5.3). The current selection of literature is by no means complete; however it might be sufficient to obtain a surface view on important principles and conclusions.

5.1 Change blindness

The change blindness phenomenon refers to the surprising difficulty observers have in noticing large changes to visual scenes [8, 30]. According to [31], the general research interest in change blindness was triggered by the discovery in the 90's that observers often failed to notice large changes to photographs that were made during an eye movement [32]. For example, 50% observers failed to notice when two cowboys sitting on a bench exchanged heads! Similar change detection failures have been found from many other experimental techniques which aim at examining the localization of the change ([33] provides an overview). Research findings explain that the observers never consciously form a complete and detailed representation of their surroundings [30], so attention is needed to select objects for processing consciously. Attention is necessary to bind different features into a complete mental representation of an object and observe changes [34]. As a result, feature changes in the attended area are more likely to be perceived and recognized. Studies on the change blindness phenomena also provide some insights into the relations between attention and visual short-term memory. As the capacity of visual short-term memory is limited, visual attention is also limited. It can be only distributed to 4 to 5 items at a time [8].

Interestingly, attention is necessary for change detection. But attention alone does not guarantee it. Changes to attended objects are still likely to go unnoticed, particularly when the changes are unexpected. Changes are reportable only once they are attended, encoded in the memory, and compared [35]. Experiments have shown that even if all the changes were attended, encoded, and remained available, change blindness might nevertheless occur due to the failure of the necessary compare process [31, 36]. Intuitively, an explanation to this might be that the working memory is fully loaded to compare and maximize the detection of one change. During this processing period, there is no extra capacity to process other possible encoded changes remaining in the visual short-term memory. As the new changes keep appearing, it is likely that some encoded changes are overwritten by new attended changes, and become eventually undetected.

5.2 Visual attention guidance

Two mechanisms are known to guide visual attention: a bottom-up, fast, primitive, and stimuli-driven mechanism that guides visual attention based on the saliency of the stimuli, i.e. the state or quality of standing out relative to neighboring items, and a second slower, top-down mechanism with variable selection criteria, which directs attention under the cognitive control. However, whether visual consciousness can be reached by either saliency-based or top-down attentional selection or by both remains controversial [37].

1) Bottom-up visual attention mechanism

In this mechanism, the visual stimuli which win the competition for saliency will automatically be attended. When an object in a visual field contains some unique features (e.g. a unique color compared to its surroundings), this object seems to pop out and captures the attention [38]. Features which can attract attention when they are unique in the scene (named as pre-attentional features) include color, shape, intensity, orientation, depth, size, and curvature [39]. In a computational model for saliency [37, 40-42], input images are processed in three independent feature channels representing color, intensity, and orientation. The culmination of processing in each channel is a map that indicates the salient locations in the image with respect to only one feature dimension. The final saliency map is a linear sum of these three sub maps. In general, the color and the intensity feature play more important role than the orientation feature in the guidance of attention [41].

2) Top-down visual attention mechanism

In contrast with the bottom-up mechanism, the top-down mechanism directs attention based on top-down knowledge, under the cognitive and volitional control. Top-down knowledge, such as the task-dependent goals [43], the contextual cues in the visual environment [44-46], and the current information in the working memory [47-51], constrain what to expect and guide the attentional processes.

Contextual cuing is a memory-based phenomenon in which previously encountered global pattern information can automatically guide attention to the location of a target in a scene [52]. Experiments on visual object detection tasks have shown that contextual information provides efficient shortcuts for the detection process [45].

It has also been illustrated that human attention varies with the nature of the task [43]. By extending the bottom-up saliency map [37] with a task-relevance map which encodes the relevance of every visual location to the current task, the combined model is able to attend at different locations, given different tasks.

The current information held in the working memory is another factor which has been proved to play a role in attention guidance. It was suggested that the item which is kept active in working memory automatically results in a tendency for attention to be attracted to stimuli that are related to that item, either visually or semantically [48]. This effect is also named as "consonance-driven orienting". In the experiment reported in [50], the subjects were asked to imagine an object and then search for a target that was interposed between some line drawings. The result suggests that people cannot help to attend to any object belonging to the same category as the object that they just imagined. Similar study described in [49] also shows that a replica of contents presented to the subject and stored in working memory seemed to draw attention even when there was no motivation for this. It was also found that semantic relations have the same attention guidance effect. The word report test in [51] shows that the word which was semantically related to the first-shown prime word tended to be reported. Study in [48] goes beyond the previous works by showing that consonance-driven orienting can play a decisive role in the control of attention.

5.3 Cross-modal attention

A great deal has been known about the visual attention. A relatively new research trend is to look at the cross-modality attention, i.e. how does attention spread among different modalities. Studies have shown that cross-modal links in spatial attention certainly exist and have influences on the reflexive or voluntary shifts of attention [53]. In other words, selective spatial attention operates cross-modally, rather than separately within each sensory modality. Previous studies suggest that shifts of attention in one modality tend to be accompanied by corresponding shifts in other modalities, in a reflexive (automatic) or voluntary (controlled) manner.

An irrelevant but salient event in one modality tends to attract covert attention towards it in other modalities, in a reflexive manner [54]. Empirical studies have proved such links for many modality combinations. For example, a salient auditory event can evidently generate rapid shifts of visual attention; a tactile event on one hand evidently generates shifts of visual and auditory attention towards that side. Only exception found is that the auditory attention does not follow the shifts of visual attention.

Cross-modal links can also direct attention in a voluntary manner, rather than a reflexive manner following a salient event. This has been demonstrated with the auditory modality and the visual modality in [55]. The subjects were given a strong expectation on the spatial location where targets in a primary modality (auditory or visual) are likely to appear. It was found that their attention tends to shift to the same location for the secondary modality (visual or auditory), as well as the primary modality, even when there was no motivation for this to happen.

In summary, spatial attention typically shifts to a common location across the modality. As a result, attention is more efficiently focused on a common location across the modalities, even though the spatial locations of these signal sources can be separated to some extent [54]. Event-related brain potential (ERP) studies have provided electrophysiological evidences that there might exist a single supra-modal attentional system that operates in coordinates of external space, independently of target modality, and controls shifts of spatial attention for all modalities [56, 57]. This finding can be somehow connected to the working memory model (see section 3). One important function of the central executive component is attention management. In this case, the central executive should be the carrier of such a supra-modal attentional system.

6 Summary of findings and conclusions

The crisis management environment is a typical high cognitive load requirement environment, due to high time urgency, high information density, high uncertainty and high risk. In this research, it is important that the design of the multimodal presentation manner avoids unnecessary cognitive load for the user and guarantees all information presented is actually perceived by the user. A psychology literature study has been carried out to obtain suggestions for such an effective and efficient presentation design. In this section, we extract and summarize the relevant findings which are going to be used in this research.

The cognitive load theory suggests that the cognitive load that is imposed by an instruction material can be considered as a summary of the intrinsic load, the extraneous load and the germane load. The intrinsic load is determined by the complexity of the instruction and the expertise of the learner. The extraneous load is the unnecessary load resulting from mainly a poor design. The germane load is induced by the learner's own mental effort to comprehend the instruction. The multimodal presentation in this research is different from an instructional material. However, they are similar in the sense that their functions are both to convey information and activate comprehension, reasoning and problem-solving. Although the crisis managers have expertise at their work, the crisis information may still impose high intrinsic load due to the high density and high uncertainty. The crisis managers are supposed to be highly attentive to their decision making tasks, therefore the germane load is also high. The cognitive load theory provides us a theoretical explanation of why crisis management is a high cognitive load demand task. Furthermore, it indicates that the design quality of the presentation does have influence on cognitive load. This research aims at achieving an effective and efficient presentation manner in order to reduce the extraneous load.

The working memory model from Baddeley suggests that working memory has separate stores for short-term maintaining visual and auditory information respectively. An information processing and attentional control system, the central executive, retrieves information from the two stores and fulfils processing for different tasks (e.g. reasoning, comprehension, and learning). The dual-coding theory from Paivio proposes that verbal and non verbal information are mentally presented and processed in two separate but interconnected systems. Representations in each system are joined by associative connections. The corresponding verbal and nonverbal information in two systems are further linked by referential connections. Eventually, all representations are bound into knowledge that can be acted upon, and retried for subsequent use. A more complete understanding of working memory can be achieved by integrating these two theories. Visual and auditory information are perceived and temporarily stored in the visual and auditory short-term memory. When they are retrieved by the central executive, they are reorganized and processed in the verbal and nonverbal systems. The associative and referential connections take place in the central executive. The capacity of working memory is limited. Consequently, the capacity of each channel/system is also limited.

The multimedia learning research from Mayer et al. is a successful example of how to apply the aforementioned three theories into the design of multimedia instructional materials. Totally nine principles have been derived from the studies on nine learning effects. These principles are design guidelines for a multimedia instruction that conveys knowledge effectively and reduces the cognitive load for the learners. Due to the similarities in the design goals, these principles can certainly provide valuable suggestions for our multimodal presentation design.

First, the coherent principle suggests not to present any irrelevant material that might distract the learning process. Due to the high information density during emergency response, it seems obvious that non-crisis related information will not be presented. However, the definition of relevancy in our case needs more consideration. As the crisis information can be very diverse, some might become irrelevant for the decision making during certain time period. In this sense, they should not be presented during this period to distract the crisis managers from working on more urgent issues. The question arising here is what presentation order is better for situation diagnose and decision making. However, it is out of the scope of this literature study. It will remain for future concern.

Second, the redundancy principle suggests not to present the same verbal information with both narration and on-screen text. We intent to obey this principle with some additional conditions. We notice that visual information has the advantage over auditory information in being permanent and therefore can be referred to repeatedly while auditory information is fleeting and difficult to retrieve once heard. If a certain verbal content remains important for a period, on-screen text or other visual presentation might be anyway necessary. In order to avoid the redundancy effect, on-screen text can appear when the narration is over, for later looking-up.

Third, the multimedia principle suggests that it is better to let both verbal and nonverbal modality to share the presentation task. The modality principle further suggests that it is better to present the verbal content in the auditory modality and the nonverbal information in the visual modality. In this case, the temporal-contiguity principle says that the corresponding visual and auditory presentation should be synchronized. When the verbal and nonverbal information both have to be presented in visual modality, the spatial-contiguity principle says that the locations of corresponding contents should be physically integrated.

Finally, the personalization principle suggests that presenting verbal information in conversational style is better than formal style. Students work harder when they feel personally involved in the conversation. This principle might not hold for the crisis management professionals with high expertise. However, it might be wise to avoid long sentences with complex grammatical structure. Keep sentences short and precise.

The findings for reducing cognitive load have been summarized above. Before applying the knowledge, we first have to guarantee that the presented information is actually perceived by the user. If the some presented information remains unnoticed, the effort on making better use of modalities to reduce cognitive load becomes useless. The research on the change blindness phenomenon concludes that attention is necessary to detect change. In the normal condition (no cognitive overload), we hope each piece of presentation does get the user's attention. According to the current knowledge about visual attention guidance, the easiest way to direct visual attention seems to be using unique features. Let the visual item that needs to be attended have certain unique feature, such as a unique color or size. Under the control of the bottom-up visual attention mechanism, this item will automatically "pop out" from its surrounding. Visual attention can be also switched by the top-down mechanism, which is mainly under the cognitive and volitional control of the observer. It seems that this mechanism can hardly be influenced by the presentation design. Furthermore, the spatial attention can be directly by cross-modality interactions. For example, an auditory signal coming from the left side may efficiently switch the user's visual attention towards the same side.

In the next research stage, the knowledge summarized above is going to be applied to the design of modality planning strategies. There is another issue which was briefly looked at during this literature study but not included in this report. That is the cognitive load measurement methods [14, 15, 58]. In the future experiment stage, we intend to access the users' cognitive load during the presentation, in order to evaluate the design. As the measurement method of cognitive load is independent from the design of modality planning strategies, it is separated from this report and will be better investigated later.

7 Acknowledgements

This research is part of the Interactive Collaborative Information System (ICIS) program (http://www.decis.nl/html/icis.html). ICIS is sponsored by the Dutch Ministry of Economic Affairs, grant nr: BSIK03024.

Reference

[1] Y. Cao, "An information assistant system for the prevention of tunnel vision in crisis management," Centre for Telematics and Information Technology, University of Twente, Enschede, the Netherlands, TR-CTIT-08-35, ISSN 1381-3625, 2008.

[2] A. D. Baddeley, "Is working memory still working," *American Psychologist,* vol. 56, pp. 851-864, 2001.

[3] G. A. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information," *Psychological Review,* vol. 63, pp. 81-97, 1956.

[4] A. D. Baddeley and G. J. Hitch, "Working memory," *The psychology of learning and motivation: advances in research and theory,* vol. 8, pp. 47-89, 1974.

[5] A. Baddeley, "Working memory: looking back and looking forward," *Nature Reviews Neuroscience,* vol. 4, pp. 829-839, 2003.

[6] A. D. Baddeley, *Essentials of Human Memory*: Psychology Pr, 1999.

[7] J. K. O'Regan, R. A. Rensink, and J. J. Clark, "Change-blindness as a result of'mudsplashes'," *Nature,* vol. 398, p. 34, 1999.

[8] D. J. Simons and D. T. Levin, "Change blindness," *Trends in Cognitive Sciences,* vol. 1, pp. 261-267, 1997.

[9] A. Paivio, *Mental Representations: A Dual Coding Approach*: Oxford University Press, USA, 1986.

[10] J. M. Clark and A. Paivio, "Dual coding theory and education," *Educational Psychology Review,* vol. 3, pp. 149-210, 1991.

[11] J. Sweller, "Cognitive load during problem solving: Effects on learning," *Cognitive Science,* vol. 12, pp. 257-285, 1988.

[12] J. Sweller, J. J. G. van Merrienboer, and F. Paas, "Cognitive Architecture and Instructional Design," *Educational Psychology Review,* vol. 10, pp. 251-296, 1998.

[13] J. Sweller, "Cognitive load theory, learning difficulty, and instructional design," *Learning and Instruction,* vol. 4, pp. 295-312, 1994.

[14] R. Brünken, J. L. Plass, and D. Leutner, "Direct measurement of cognitive load in multimedia learning," *Educational Psychologist,* vol. 38, pp. 53–61, 2003.

[15] R. Brünken, J. L. Plass, and D. Leutner, "Assessment of Cognitive Load in Multimedia Learning with Dual-Task Methodology: Auditory Load and Modality Effects," *Instructional Science,* vol. 32, pp. 115-132, 2004.

[16] P. Chandler and J. Sweller, "Cognitive Load Theory and the Format of Instruction," *Cognition and Instruction,* vol. 8, pp. 293-332, 1991.

[17] P. Chandler and J. Sweller, "Cognitive Load While Learning to Use a Computer Program," *Applied Cognitive Psychology,* vol. 10, pp. 151-170, 1996.
[18] S. Kalyuga, P. Chandler, and J. Sweller, "Managing Split-attention and Redundancy in Multimedia Instruction," *Applied Cognitive Psychology,* vol. 13, pp. 351-371, 1999.
[19] S. Y. Mousavi, R. Low, and J. Sweller, "Reducing Cognitive Load by Mixing Auditory and Visual Presentation Modes," *Journal of Educational Psychology,* vol. 87, pp. 319-34, 1995.
[20] R. E. Mayer, "Multimedia Learning: Are We Asking the Right Questions?," *Educational Psychologist,* vol. 32, pp. 1-19, 1997.
[21] R. E. Mayer, "Animation as an Aid to Multimedia Learning," *Educational Psychology Review,* vol. 14, p. 87, 2002.
[22] R. E. Mayer, "The promise of multimedia learning: using the same instructional design methods across different media," *Learning and Instruction,* vol. 13, pp. 125-139, 2003.
[23] R. E. Mayer, "Elements of a Science of E-Learning," *Journal of Educational Computing Research,* vol. 29, pp. 297-313, 2003.
[24] R. E. Mayer, J. Heiser, and S. Lonn, "Cognitive Constraints on Multimedia Learning: When Presenting More Material Results in Less Understanding," *Journal of Educational Psychology,* vol. 93, pp. 187-98, 2001.
[25] R. E. Mayer and R. Moreno, "A Split-Attention Effect in Multimedia Learning: Evidence for Dual Processing Systems in Working Memory," *Journal of Educational Psychology,* vol. 90, pp. 312-320, 1998.
[26] R. E. Mayer and R. Moreno, "Nine Ways to Reduce Cognitive Load in Multimedia Learning," *Educational Psychologist,* vol. 38, pp. 45-52, 2003.
[27] R. E. Mayer and V. K. Sims, "For Whom Is a Picture Worth a Thousand Words? Extensions of a Dual-Coding Theory of Multimedia Learning," *Journal of Educational Psychology,* vol. 86, pp. 389-401, 1994.
[28] R. Moreno and R. E. Mayer, "Cognitive principles of multimedia learning: The role of modality and contiguity," *Journal of Educational Psychology,* vol. 91, pp. 358-368, 1999.
[29] R. E. Mayer and R. Moreno, "Aids to computer-based multimedia learning," *Learning and Instruction,* vol. 12, pp. 107-119, 2002.
[30] R. A. Rensink, J. K. O'Regan, and J. J. Clark, "The Need for Attention to Perceive Changes in Scenes," *Psychological Science,* vol. 8, pp. 368-373, 1997.
[31] D. J. Simons and R. A. Rensink, "Change blindness: past, present, and future," *Trends in Cognitive Sciences,* vol. 9, pp. 16-20, 2005.
[32] J. Grimes, "Scenes across Saccades," *Perception,* vol. 5, pp. 89-110, 1996.
[33] R. A. Rensink, "Change detection," *Annual Reviews in Psychology,* vol. 53, pp. 245-277, 2002.
[34] A. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognitive Psychology,* vol. 12, pp. 97-136, 1980.
[35] M. E. Silverman and A. Mack, "Change blindness and priming: When it does and does not occur," *Consciousness and Cognition,* vol. 15, pp. 409-422, 2006.

[36] S. R. Mitroff, D. J. Simons, and D. T. Levin, "Nothing compares 2 views: Change blindness can occur despite preserved access to the changed information," *Perception & Psychophysics,* vol. 66, pp. 1268-1281, 2004.
[37] L. Ltti and C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention," *Vision Research,* vol. 40, pp. 1489-1506, 2000.
[38] E. Sussman, I. Winkler, and E. Schröger, "Top-down control over involuntary attention switching in the auditory modality," *Psychonomic Bulletin & Review,* vol. 10, pp. 630-637, 2003.
[39] J. M. Wolfe, "Visual attention," *Seeing,* vol. 2, pp. 335-386, 2000.
[40] R. J. Snowden, "Visual Attention to Color: Parvocellular Guidance of Attentional Resources?," *Psychological Science,* vol. 13, pp. 180-184, 2002.
[41] D. Parkhurst, K. Law, and E. Niebur, "Modeling the role of salience in the allocation of overt visual attention," *Vision Research,* vol. 42, pp. 107-123, 2002.
[42] L. Ltti and C. Koch, "Computational modelling of visual attention," *Nature Reviews Neuroscience,* vol. 2, pp. 194-203, 2001.
[43] V. Navalpakkam and L. Itti, "A goal oriented attention guidance model," *Lecture notes in computer science,* vol. 2525, pp. 453-461, 2002.
[44] M. M. Chun and Y. Jiang, "Top-Down Attentional Guidance Based On Implicit Learning of Visual Covariation," *Psychological Science,* vol. 10, pp. 360-365, 1999.
[45] A. Oliva, A. Torralba, M. S. Castelhano, and J. M. Henderson, "Top-down control of visual attention in object detection," *Image Processing, 2003. Proceedings. 2003 International Conference on,* vol. 1, 2003.
[46] M. S. Peterson and A. F. Kramer, "Attentional guidance of the eyes by contextual information and abrupt onsets," *Perception & Psychophysics,* vol. 63, pp. 1239-1249, 2001.
[47] R. Desimone, "Neural mechanisms for visual memory and their role in attention," *Proceedings of the National Academy of Sciences,* vol. 93, p. 13494, 1996.
[48] L. Huang and H. Pashler, "Working memory and the guidance of visual attention: Consonance-driven orienting," *Psychonomic Bulletin & Review,* vol. 14, pp. 148-153, 2007.
[49] P. E. Downing, "Interactions Between Visual Working Memory and Selective Attention," *Psychological Science,* vol. 11, pp. 467-473, 2000.
[50] H. Pashler and L. P. Shiu, "Do images involuntarily trigger search? A test of Pillsbury's hypothesis," *Psychon Bull Rev,* vol. 6, pp. 445-8, 1999.
[51] V. J. Dark, K. G. Vochatzer, and B. A. Van Voorhis, "Semantic and spatial components of selective attention," *J Exp Psychol Hum Percept Perform,* vol. 22, pp. 63-81, 1996.
[52] M. M. Chun and Y. Jiang, "Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention," *Cognitive Psychology,* vol. 36, pp. 28-71, 1998.
[53] J. Driver and C. Spence, "Crossmodal attention," *Current Opinion in Neurobiology,* vol. 8, pp. 245-253, 1998.
[54] J. Driver and C. Spence, "Cross-modal links in spatial attention," *Philosophical Transactions of the Royal Society B: Biological Sciences,* vol. 353, p. 1319, 1998.

[55] C. Spence and J. Driver, "Audiovisual links in endogenous covert spatial attention," *J Exp Psychol Hum Percept Perform,* vol. 22, pp. 1005-30, 1996.
[56] M. Eimer, "Can attention be directed to opposite locations in different modalities? An ERP study," *Clinical Neurophysiology,* vol. 110, pp. 1252-1259, 1999.
[57] M. Eimer, J. van Velzen, B. Forster, and J. Driver, "Shifts of attention in light and in darkness: an ERP study of supramodal attentional control and crossmodal links in spatial attention," *Cognitive Brain Research,* vol. 15, pp. 308-323, 2003.
[58] F. Paas, J. E. Tuovinen, H. Tabbers, and P. W. M. Van Gerven, "Cognitive Load Measurement as a Means to Advance Cognitive Load Theory," *Cognitive Load Theory: A Special Issue of Educational Psychologist,* 2003.