

SOUND AS AFFECTIVE DESIGN FEATURE IN MULTIMEDIA LEARNING – BENEFITS AND DRAWBACKS FROM A COGNITIVE LOAD THEORY PERSPECTIVE

Anke Königshulte

University of Applied Sciences Bremerhaven, Germany

ABSTRACT

The study presented in this paper investigates the potential effects of including non-speech audio such as sound effects into multimedia-based instruction taking into account Sweller's cognitive load theory (Sweller, 2005) and applied frameworks such as the cognitive theory of multimedia learning (Mayer, 2005) and the cognitive affective theory of learning with media (Moreno, 2006). Proceeding from the assumption that sound is an incisive means to affect people's emotional state it is argued that sound may also be well suited to stimulate involvement and motivation in learning situations, thereby bringing the learner to invest more mental effort into learning, which finally leads to better learning performance.

This paper refers to an experimental case study, which was carried out within the framework of a Master's Thesis at the University of Applied Sciences Bremerhaven (Germany). In order to investigate the cognitive effects of including sound into multimedia learning, two groups of 1st semester Digital Media students were asked to learn about a historic subject using two different experimental designs: One version of a prototypical learning application consists of a photo slideshow with accompanying audio narration and another version consists of the same material supplemented with environmental sounds that illustrate the content of the lesson.

Comparing both groups, the results don't reveal significant differences in learning performance. However, the subjective mental effort ratings of the participants are identified as a positive predictor for the performance score and are thus hypothetically discussed as being an indicator for learner motivation. The analysis finally confirms that the learner involvement, which is a measure relating the performance score and the mental effort ratings (Paas et al., 2005), during the subsequent achievement test is significantly higher when sounds were presented during instruction. These results suggest that the inclusion of sound may have positive effects on motivation and learning, which is according to a cognitive-motivational theory.

KEYWORDS

Multimedia Learning, Instructional Design, Sound, Cognitive Load, Motivation, Involvement

1. INTRODUCTION

Through the introduction of digital audio in the early 1980s and the evolution of Internet capabilities, appropriate and cheaper software and (portable) audio players, multimedia developers are nowadays able to easily store, manipulate, (re-)use, and thus fully integrate sound in (instructional) software. However, even though there have been significant changes in technology since the heyday of radio in the 1930s, only "little has been done to advance the single notion of audio in instruction", Holmes and LaBoone (2002) argue (p. 57, cited in Schlosser & Burmeister, 2006).

Research within the realm of auditory perception (e.g. McAdams & Bigand, 1993; Van Leeuwen, 1999) as well as common practice in film sound (e.g. Chion, 1994) and computer game sound (e.g. Jørgensen, 2009) suggests that sound offers many promising possibilities in instructional design that remain largely unexplored. Without a doubt, auditory information participates fundamentally in the development of knowledge by facilitating the acquisition, processing, and retrieval of information in many ways (McAdams & Bigand, 1993, p. 1-9; Gaver, 1993). However, the integration of non-speech audio, that is sound and music in its different forms, has been apparently neglected in the current state of research and practice in

multimedia learning. Although some applications indeed add sound (because it is possible), there are very few research studies that formally evaluate the effects of it, respectively, provide guidelines or a structured way to integrate sound in instructional media (Bishop et al., 2008).

1.1 Objective and Framework

Sound is a very different medium for representing information than visual images. Due to natural factors, such as the absence of some mechanism to shut the ears (“eyelids”), and the omnidirectionality of hearing, listening is much less limited than seeing. Therefore, people encounter sounds and noises every time and everywhere, so that they usually ‘fade them out’ as long as it is not relevant to (re-)act to them. In everyday listening, people naturally seek to denote sounds to their sources, actions, and events that cause them in order to get important details about the immediate surrounds; it is, like Smalley (1996, p.79) puts it, “a question of living and acting in the world, ultimately of survival”. When a person walks along a road at night hearing a car he/she most likely does not focus on the sound itself, e.g. its pitch and loudness, but compares the sound to his/her memories for the known objects that make that sound. With attending to the whole scene, drawing from existing schemas and experiences the person will understand the situation, e.g. it is a car with a strong engine approaching quickly from behind and I should step out of the path (Gaver, 1993, p.1). It may be for these reasons that sound, even not alarming by itself such as an even faraway motor’s engine or flat tire’s faint thumping, can immediately activate existing schemas and is thus generally more effective than images for gaining attention (Posner et al., 1976; Bernstein & Edelstein, 1971, cited in Bishop & Cates, 2001, p.11). Sounds, which also carry strong affective significance, may add yet another dimension to the audiovisual by making the environment more immersive, tangible and by provoking emotions, such as suspense. That way, even unintrusive sounds, as for example waves hitting the shore or whistling tree leaves, may also effectively hold attention over time and reduce distraction of competing stimuli (Thomas & Johnston, 1984, cited in Bishop & Cates, 2001, p.12). Chion (1994) points out, that it is the sound’s physical nature as vibrant phenomenon and its omnipresence that interferes with and affects people’s perception more than the image, especially when people do not give conscious attention to it. Therefore, sound can be an insidious means of affective and semantic manipulation (pp.33). Sound facilitate a rather intuitive view of the information it presents than the visual.

The present study draws mainly on the emotive qualities of sound and tries to approach in how far the inclusion of sound (i.e. non-speech audio) to audiovisual learning material can affect the learners’ affective state; in particular, if it can stimulate learners’ involvement and motivation to learn in a way that learning can be enhanced.

Nevertheless, numerous of experimental media studies reveal that the actual challenge for the instructional multimedia design is for the processing demands to not exceed the limited capacity of the cognitive system, which otherwise easily end up in creating a cognitive overload and decreased learning performance (e.g. Moreno & Mayer, 2000; Mayer et al., 2001). Moreno and Mayer (2000, p.118) accordingly conclude that any additional material “that is not necessary to make the lesson intelligible or that is not integrated with the rest of the materials will reduce effective working memory capacity and thereby interfere with the learning of the core material”. It is thus especially a cognitive perspective on learning that raises doubts about the effectiveness of including sound in instruction. Accordingly, Bishop and Cates (2001, p.6) argue that “without a strong theoretical cognitive foundation, the sounds used in instructional software may not only fail to enhance learning, they may actually detract from it“. In line with this, this work aims to integrate and design sound with regard to research-based theory of how students learn in order to successfully enhance learning.

1.2 Theory and Predictions

Cognitive theories of multimedia learning, such as Sweller’s *cognitive load theory* (Sweller, 1988; Sweller, 2005), and Mayer’s *cognitive theory of multimedia learning* (Mayer, 2005), clarify memory processes of learning and problem solving and have thus contributed a solid foundation to guidelines that help practitioners to design multimedia instruction more effectively. According to the *cognitive theory of multimedia learning* (CTML, see Figure 1), different stages of memory process audio/verbal and visual/pictorial information in distinct, independent but cooperative channels. Meaningful learning only

occurs through active processing in working memory, that is, selecting, organizing and integrating information with the aim to construct coherent mental representations. These processes are partly guided by prior knowledge stored in long-term memory.

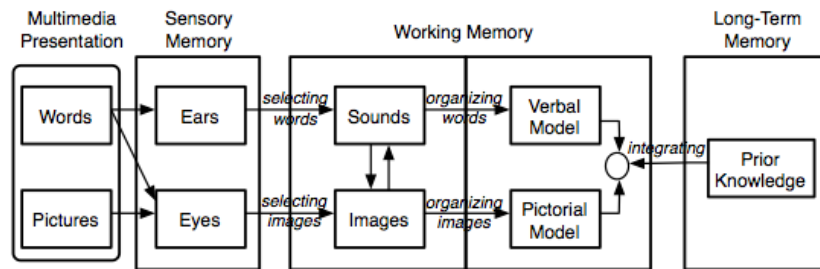


Figure 1. CTML, according to Mayer (2005)

The working memory processes only a very limited amount of information at any one time, which is a major impediment when learners are required to learn new material. Therefore, learning instruction should avoid to provoke processing and storage that is not relevant for learning, which is the primary objective set by Sweller's *cognitive load theory*. Cognitive load represents the load that performing a particular task imposes on the learner's cognitive system and is determined by the interaction of environment, learner, and task characteristics. On this account, three types of cognitive load are distinguished (see Sweller et al., 1998, pp.259): The intrinsic load is the inherent difficulty level of the topic to be learned, extraneous load is considered as unnecessary processing and is determined by the manner in which information is presented, and germane load is devoted to learning, that is, processing, construction, and automation of schemas. Cognitive load theory and CTML focus on instructional methods aimed at reducing extraneous load and that way freeing memory capacity for an increase in germane learning activities (Sweller, 2006, p.168). Several empirically validated methods for eliminating extraneous load via the instructional design have been published yet, such as the modality principle, redundancy principle, coherence principle, signaling, segmentation, and many more (e.g. Mayer, 2005; Mayer & Moreno, 2003).

In Moreno and Mayer's (2000) study investigating the inclusion of seductive details in multimedia instruction, participants perform significantly worse in subsequent retention and transfer tests when they have learned on basis of narrated animations with looped bland background music or with sound effects. With reference to CTML, the authors suggest that sound and music constitutes unnecessary (extraneous) load that might prime the activation of inappropriate prior knowledge as the organizing schema in working memory. Furthermore, additional sounds risk to reduce the learners' capacity for integrating the relevant verbal and visual material to a coherent system, which impedes learning due to memory's capacity limitations.

Cognitive load theory and CTML, however, focus mainly on freeing working memory capacity by reducing extraneous load, which assumes that learners would automatically spend all their available resources in productive ways, that is, in germane learning activities. Accordingly, Paas et al. (2005, p.25) criticizes: "Until now, cognitive load theory (CLT) has focused on the alignment of instruction with cognitive processes, without recognizing the role of motivation in training". It is important to note though, that learners' investment of cognitive resources can be rather affected by relatively instable factors such as general orientation, motivation and state of arousal. The *cognitive affective theory of learning with media* (CATLM) extends CTML to better integrate the role of motivation, that is, affect and motivation that mediate learning by increasing or decreasing the amount of cognitive resources that learner invests on the task (Moreno, 2006; Moreno & Mayer, 2007). On that account, different germane load inducing techniques have been investigated in recent studies. These reveal that, for instance, emotional design features (e.g. Plass et al., 2014), interesting additional material (e.g. Park et al., 2011) and methods such as imagination assignments (e.g. Cooper, et al., 2001) may lead to an increase in mental effort investment and thereby learning performance.

These considerations form the theoretical case for including sound into multimedia instruction presuming that the addition of sound may be a useful tool to direct attention, add emotion and "playfulness" to the lesson and encourage imagination. According to CATLM, it is hypothesized that sound in multimedia lessons can stimulate learner's interest and enjoyment and achieve a greater level of motivation to engage in deeper learning. Sound as affective design feature thus encourages the learner to invest more mental effort into learning, which lead to a better learning performance.

2. EXPERIMENT

The aim of the experiment is to investigate whether the addition of corresponding environmental sounds to a narrated photo slideshow increases the learner's involvement and motivation, thereby encourages to invest more mental effort into learning, which, according to CATLM, leads to better learning performance.

Two groups of students are asked to learn using two different experimental designs, which are with or without sound illustrations. Via subsequent achievement tests and questionnaires the performance score, invested mental effort, motivation, and user experience related data is measured and compared using quantitative methods.

2.1 Learning Material

For the experimental test, a learning application is prototypically implemented that teaches the user about the history and (re-)construction of the medieval cog found in Bremen, which is exhibited in the German Maritime Museum Bremerhaven. The learning application, which is installed on desktop computers in the computer lab of the University, is segmented into five screens, each one presents a slideshow lasting about three minutes including two or three pictures and accompanying audio explanations. The learner can use the standard media control buttons *play/pause*, *rewind*, *repeat*, *volume* and can move to the next or previous screen using the mouse when they want to.

One version of the application includes environmental sounds, which adapt during runtime according to the content of the lesson. As the sounds are included with the aim to attain positive effects on learning efficiency, the requirements is set to minimize the (feared) likelihood of cognitive overload and distraction. Considering the theoretical considerations summarized above, it is important with the sounds to not just add "appealing" extraneous load. As auditory information can be far more attention-grabbing than visual pictures, only relevant, not too intrusive sounds are applied deliberately and rather sparingly in order to not distract from the other material. The sound needs to fit in with the context and is arranged in a one-to-one correspondence with the related spoken text (or "idea unit" of the text) and pictures so that the information can be easily associated with each other. Even spoken text and other sounds can be perceived when presented simultaneously, this work suggests to present verbal and related non-verbal auditory information somewhat successively due to possible interferences. For the sound to be intelligible and meaningful, the application presents them with and shortly after the related text unit and/or picture. That way, text or picture, which are certainly more self-contained and self-explanatory, provide the necessary context to enable a clear attribution of the sound.

In order to verify these assumptions and design decisions, one "control chapter" of the application does not comply with the design principles as made above. Here, rather ambiguous and intrusive sounds are played simultaneously to the spoken text and alongside animated pictures, which is supposed to make cognitive load higher, holding the risk of reducing the learners' capacity for processing relevant information.

Even though sound can take numerous different forms, such as speech, alarm signals and "beep"-sounds, music, and more, considerations in this experiment are limited to the inclusion of environmental sounds with the aim to facilitate imagery and a dual coding of information (Paivio, 1986; Thompson & Paivio, 1994). That means, environmental sounds included are referred to real events related to the content of the lesson and have meaning based on their causal relation to that event. For example, the picture of a cog (replica) at sea and its general information is supported by the sound of sailing, that is mainly wind and waves plus a clear deep wooden creaking and the flapping of a sail, in order to increase the listeners immersion into the content and to make him/her feel the dangers of such journeys. These sounds should mentally refer to and reinforce the sound-producing source, which is the cog as wooden – sailing – ship depicted by the picture. Further sounds illustrate and strongly highlight the circumstance the cog sank. Hammering on wood (it sank when still being under construction), the sound of whistling wind, storm, flood, creaking and breaking wood, and crashing depicts the process of the sinking cog during a storm. These sounds are all source-disconnected ("off camera"), that is, neither ship-building nor the sinking cog is depicted by the pictures, and should thus increase the listener's imagination. The use of the capstan is illustrated by the sounds of using the capstan, quarter turnings, wooden tool, sticks and quarter turnings, panting of men, the latter "off-camera", which intends to make this procedure rather palpable and reinforce the information that this tool was made of wood, heavy to use – hard work – and needed to be used by multiple man.

2.2 Method

2.2.1 Measures and Instruments

The *prior knowledge* is controlled via a short pre-knowledge questionnaire before instruction. Five open questions assess the participants prior knowledge and understanding of basic concepts concerning the Hanseatic League and cogs in general.

Learning performance is gained from learning objectives achieved by the participants, which is measured by points given for correctness and/or accuracy of answers. An achievement questionnaire that has to be completed directly after instruction tests memory of the material via retention questions (mostly multiple choice) and deeper understanding via open and transfer questions.

Mental effort refers to the cognitive capacity that is allocated to accommodate the demands of the task. Due to practical reasons, data is collected through self-reports of the participants using a 7-point likert-scale, which is a valid and widely-used tool for measuring mental load as supposed to correlate highly with objective measures (e.g. Kalyuga et al. 2000, p.130). Participants are asked to rate the mental effort they needed for the task by translating it into numerical values from 1 (very low) to 7 (very high) directly after instruction and after the achievement-test, so that the values indicate mental effort associated with learning and with the performance test.

Task involvement combines measures of mental effort and performance. Presuming that motivation, mental effort invested on the task, and performance are positively related, Paas et al. (2005) theorizes that the combination of mental effort and performance score is a more accurate measure of motivational effects of instructional conditions than subjective rating scales. High performance associated with high effort is called high involvement instruction, whereas low-task performance with low effort is called low involvement instruction. The computational model of task involvement (respectively learners' motivation on the task) calculates the relative involvement of learners by mapping z-values of mental effort and performance score using a particular formula as derived and depicted in Paas et al. (2005, p. 29).

The "*User Experience Questionnaire*" ("UEQ"¹), which allows the participants to express feelings, impressions, and attitudes towards the application, is utilized to gain further quantitative data that helps to interpret (as supposed to mediate) the performance results. In particular, the UEQ is supposed to measure the affective and attentional factors the sound might have an influence on. This test uses the scales *Attractiveness*, which assesses general impressions towards the application, and *Stimulation*, which assesses if the application is interesting and exciting to use and if the learner feels motivated to further use it.

2.2.2 Participants and Procedure

The participants are 38 Digital Media Bachelor students in their 1st semester recruited from the University of Applied Sciences in Bremerhaven (Germany). The mean age is 22 years ranging from 18 years to 32. 42% of the sample are male students.

Two different versions of the multimedia-based learning module "A Guided Tour of the Hanseatic Cog" are copied onto computers in the computer lab before the experiment. The participants are randomly assigned to one of the two experimental groups:

(1) One group of participants receive a slideshow containing pictures and verbal information in form of audio narration (basic-group).

(2) The same material as in (1) but additionally supported by corresponding environmental sounds that illustrate the content of the lesson (sound-group).

Participants are not informed about the research question and the differences of the two versions of learning material. The test is conducted in three sessions with 11 – 13 participants each lasting around 75 minutes. Each student completed the instruction individually using a personal computer and good quality headphones, which lasts about 20 minutes. Once the instructional program is finished each participant get the first post-test questionnaire, which asks the participants to rate their invested mental effort during instruction, followed by the 13 items retention and transfer questions. This questionnaire concludes again asking for the invested mental effort on the achievement test. After that, the participants are asked to fill out a second form including further two transfer questions and the UEQ.

¹ <http://www.ueq-online.org/>; (perceived July, 2015)

2.3 Results

2.3.1 Preliminary Study

Most of the participants have very little prior knowledge about the topic Hanseatic League and cogs as indicated by the pre-test questionnaire. However, four participants are excluded from further calculation because of high scores in the pre-knowledge test. A homogeneous sample of 34 participants remains for this study; having 16 students in the basic-group (group 1), 18 in the sound-group (group 2). Simple t-tests indicate that finally both groups do not differ significantly (with $p < .05$) in basic characteristics, that is, pre-knowledge, gender, mean age, and initial motivation regarding the test and the learning topic.

The reliabilities for the post-test scales are calculated using Cronbach's Alpha, which reveals satisfactory results for the 17 items of the post-achievement test ($\alpha=0,78$).

2.3.2 Learning Performance

As Figure 2 shows, the sound-group's performance score is higher on the descriptive level in all types of question, i.e. open, retention, and transfer questions. However, separate analyses of covariance (ANCOVA) with the experimental condition as independent variable, the performance scores as dependent variables and the pre-knowledge as covariate (as there is a strong positive correlation between performance and pre-knowledge, with $r=.56$; $p=.00$), shows that the sound-group did not perform significantly better than the basic-group on the retention nor on the transfer part of the post-test.

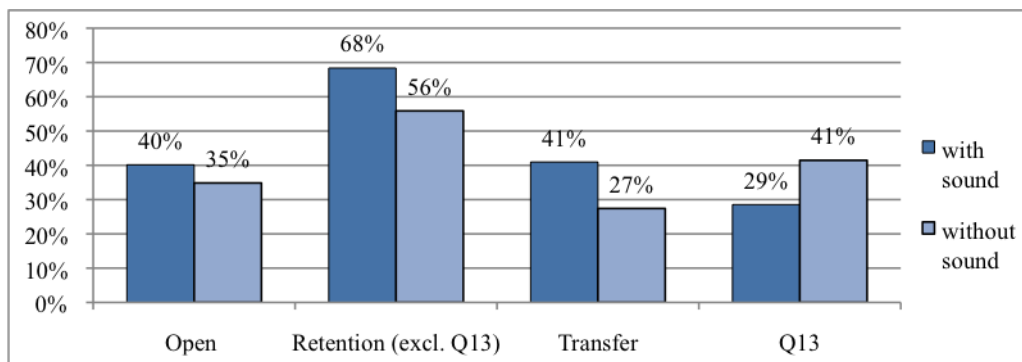


Figure 2. Performance score

In order to verify the design decisions theoretically set in order to prevent cognitive overload, the control chapter, which is covered by question 13, is evaluated separately. An analysis of the score of the nine retention questions show that the results of question 13a and 13b exceptionally differs from the other results in a way that the basic-group considerably outperformed the sound-group (Figure 2).

Overall, there is no evidence that the inclusion of sound within audiovisual learning material improves transfer and retention performance in learning. However, the results suggest that learning performance depends on its design, such as its intrusiveness and arbitrariness, and its integration with the rest of the material in a way that sound may impede retention if it is not integrated deliberately.

2.3.3 Mental Effort

For the mental effort ratings separate analyses of variance (ANOVA) were conducted using the experimental condition as independent variable and the mental effort ratings for the instruction and for the test phase as dependent variables. Table 1 shows that the mean of mental effort, which are mapped to a scale from -3 to +3, are higher for the sound-group during the test phase, but do not differ significantly.

Table 1 Mental effort ratings mapped to a scale from -3 to +3

	mental effort ratings test phase					mental effort ratings learning phase				
	N	M	SD	F	p	N	M	SD	F	p
1) without sound	16	,66	1,01	4,03	,12	16	,25	,68	1,14	,43
2) with sound	18	1,11	,63			18	0	1,08		

2.3.4 Learner Involvement

A regression calculation indicate that mental effort during the test phase may be a positive predictor for the performance score (with $p=.04$), and thus supports the cognitive-motivational viewpoint as depicted by the learner involvement construct. Accordingly, the z-values of mental effort ratings and performance score of the transfer questions – as being an indicator for germane load investment and “deep” learning – are mapped on a Cartesian axis in order to calculate the task involvement of the learner. Figure 3 shows that the sound group is clearly in the high involvement area compared to the basic group.

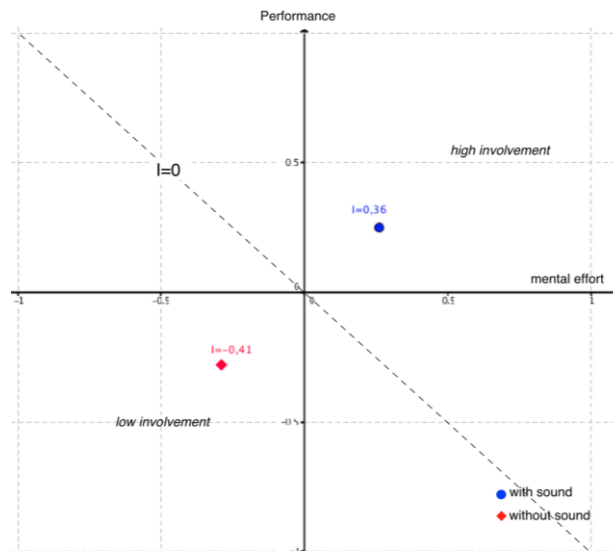


Figure 3. Learner involvement during test phase

An ANOVA reveals that the sound-group ($M=0,36$; $SD=0,92$) shows a significant higher learner involvement during the achievement test than the basic-group ($M=-0,41$; $SD=1,25$), with $p < .05$.

2.3.5 User Experience

Even though the sound-group is expected to state higher affective ratings than those of the basic-group, the groups did not differ significantly on their mean ratings of attractiveness or stimulation of the learning material.

3. CONCLUSION

3.1 Summary of Main Findings

The study does not confirm differences in retention and transfer performance nor in subjective mental effort ratings during learning, which suggests that sound, as designed, arranged and coordinated in this work, just do not impose additional cognitive load during learning nor have other effects on cognitive processes, which benefit or hinder learning. Another possible interpretation is that the advantages concerning motivation (as assumed in this study) can be offset by the distraction disadvantages of the environmental sounds, which results in no differences in the learning performance between learners that have learned with audio-enhanced narratives and the basic version without sound.

However, a regression calculation confirms that the invested mental effort in the achievement test is a predictor for the performance score, that is, the more mental effort the participants invested in answering the test questions the better they scored in the test. As a result of the higher performance score and mental effort ratings, the study finally shows a significant higher learner involvement of the sound-group, which supports the notion of sound having a positive influence on the learner’s motivation to engage in appropriate learning

processes. However, as there is no difference in mental effort investment during learning, the results suggest that motivational effects may exert its influence only during the subsequent performance test.

3.2 Discussion

The study does not identify the mechanisms by which the integration of sound encourage learners to invest more effort in answering the test questions. Against expectation, the results reject that the higher involvement of the sound-group is attributable to the audio-enhanced narratives being more stimulating and appealing, as the UEQ scales reveal. One possible interpretation of the results, even though rather speculative, refers to positive imagination effects induced by the environmental sound. The instructional material may have encouraged imagery and concreteness of the learning topic and to process the concept of the material actively and more deeply, which could have lead to an increase in mental effort investment conducive to learning. This notion is supported by participants' statements that the presentation of sound has made them feel rather immersed into the content of the lesson. However, this could be an interesting topic for further investigations, which should also include appropriate qualitative measures. It has to be noted that the use of subjective rating scales and with it the participants ability to reliably value his/her affective state and invested mental effort has clearly its limits.

It can be concluded that the results do not confirm Moreno and Mayer's (2000) study that reveal detrimental learning effects induced by the incorporation of sound effects into audiovisual learning material. The deliberate design and inclusion of relevant sound as it is done in the application used in this experiment, does not increase extraneous cognitive load nor has negative effects on learning performance. However, practically, the results suggest that sound also risks interfering with the processing of the otherwise essential verbal information depending on its design and integration. Therefore, the study concludes in accordance with Moreno and Mayer (2000, p.124): More studies that examine the design of sounds, its integration and coordination in multimedia lessons seem to be necessary to get solid results.

3.3 Limitations

The study can be seen as a first attempt to investigate the cognitive effects of having sound effects in multimedia instruction. It was carried out within the framework of a Master's Thesis and has thus clearly its limits – starting from the small sample size of 38 participants.

It can be assumed that only learners with a basic (at least) medium prior knowledge may profit in general from additional environmental sounds. However, the role of prior knowledge (as well as other student characteristics, such as spatial ability, age, gender, etc.) has to be subjected to further investigations as it couldn't be provided by this experiment.

The results and conclusions are also practically very limited as the study only deals with two different instructional formats, that is multimedia explanations with or without sound, and with one kind of learner. It also entirely neglects to look into different instructional methods, which should be closely connected to the choice of media and its format as well as learner characteristics (see e.g. Moreno, 2006). To extend the external validity, more research is needed in order to generalize the results to other learning environments, also more authentic ones, such as a museum or mobile guides in different contexts, also to different instructional methods (e.g. game-based methods), to other learners, as well as to other learning topics.

With the aim to get clear and solid results, the case study has concentrated on the design and evaluation of mainly one aspect of how sound can be utilized in multimedia learning, which is the role of motivation in order to improve memory performance. However, sound promises to have many more possibilities to support multimedia learning in the same way as sound can take numerous other forms than environmental sounds, such as musical sounds or other abstract sounds, which are worth investigating. For example, guiding musical themes, leitmotifs, could be a means to orientation and structuring in lessons, as they can be applied to repeating events. It can furthermore be assumed that auditory features being distinct from the content of the lesson could be also used to elicit and maintain attention and support structuring and selection of the content to be further processed in working memory. More on sound's potential role in instructional software can be found in Bishop and Cates (2001).

REFERENCES

- Bishop, M.J. et al, 2008. Sound's use in instructional software to enhance learning: A theory to practice content analysis. *Educational Technology Research & Development*, 56, pp. 467-486.
- Bishop, M.J. and Cates, W.M., 2001. Theoretical foundations for sound's use in multimedia instruction to enhance learning. *Educational Technology Research and Development*, 49 (3), pp. 5-22.
- Chion, M., 1994. *Audio-Vision: Sound on Screen*. Columbia University Press, New York.
- Cooper, G. et al, 2001. Learning by imagining. *Journal of Experimental Psychology: Applied*, Vol 7(1), pp. 68-82.
- Gaver, W., 1993. What in the world do we hear? An ecological approach to auditory source perception. *Ecological Psychology*, 5, pp. 1-29.
- Holmes, G. and LaBoone, E., 2002. The importance of culture when creating audio-enhanced, Web-based instruction. *TechTrends*, 46(2), pp. 56-61.
- Jørgensen, K., 2009. *A Comprehensive Study of Sound in Computer Games: How Audio Affects Player Action*. PhD. Mellen Press.
- Kalyuga, Set al. 2000. Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, Vol. 92, pp.126-136.
- Mayer, R., 2005. Cognitive theory of multimedia learning. In R.E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press, Cambridge, pp. 31-48.
- Mayer, R. E. et al, 2001. Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, 93, pp. 187-198.
- Mayer, R.E. and Moreno, R., 2003. Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist* 2003, 38(1), pp. 43-52.
- McAdams, S. and Bigand, E., 1993. *Thinking in sound: The cognitive psychology of human audition*. Oxford University Press, Oxford.
- Moreno, R., 2006. Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning*, Vol. 22, No. 3, pp. 149-158.
- Moreno, R. and Mayer, R.E., 2000. A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages, *Journal of Educational Psychology*, Vol. 92(1), pp. 117-125.
- Moreno, R., and Mayer, R., 2007. Interactive multimodal learning environments: Special issue on interactive learning environments: Contemporary issues and trends. *Educational Psychology Review*, Vol 19, pp. 309-326.
- Smalley, D., 1996. The Listening Imagination: Listening in the Electroacoustic Era. Contemporary, *Music Review*, 13 I(2), pp. 77- 107.
- Sweller, J., 1988. Cognitive load during problem solving: Effects on learning. *Cognitive Science* Vol. 12 (2): pp. 257–285.
- Sweller, J., 2006. The worked-example effect and human cognition. *Learning and Instruction*, Vol. 16(2), pp. 165 - 169.
- Sweller, J., 2005. Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. Cambridge University Press, pp. 19-30
- Sweller, J., et al, 1998. Cognitive architecture and instructional design. *Educational Psychology Review*, Vol. 10, pp. 251– 296.
- Paas, J. E. et al, 2005. A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction, *Educational Technology Research and Development*, Vol. 53, pp. 25–34.
- Paivio, A., 1986. *Mental Representations: A Dual Coding Approach*. Oxford University Press, New York
- Park, B. et al, 2011. Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior*, Vol. 27(1), pp. 5–10.
- Plass, J. L., et al, 2014. Emotional design in multimedia learning: effects of shape and colour on affect and learning. *Learning and Instruction*, Vol. 29, pp. 128 – 140.
- Posner, et al, 1976. Visual dominance: An information-processing account of its origins and significance. *Psychological Review*, Vol 83, pp. 157–171.
- Schlosser, C. A. and Burmeister, M. L., 2006. Audio in online courses: Beyond podcasting. http://www.nova.edu/~burmeist/audio_online.html [Feb 2011]
- Thompson, V. and Paivio, A., 1994. Memory for pictures and sounds: Independence of auditory and visual codes. *Canadian Journal of Experimental Psychology*, Vol. 48, pp. 380-398.
- Thomas, F. and Johnston, R., 1984. The Disney sounds. In W. Rawls (Ed.), *Disney animation: The illusion of life*. New York: Abbeville, pp. 145–161.
- Van Leeuwen, T., 1999. *Speech, Music, Sound*. MacMillan Press, London.