We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



5

An Empirical Investigation into the Use of Multimodal E-Learning Interfaces

Marwan Alseid and Dimitrios Rigas
University of Bradford
United Kingdom

1. Introduction

Recently, most of e-learning user interfaces mainly depend on text with graphics in the delivery of information, and frequently disregard other communication metaphors within the visual and auditory channels. The lack of multimodality results in overloading user' visual channel (Brewster 1997; Brewster 1998). Also, in some cases important information being communicated could be missed (Oakley, McGee et al. 2000). Including more than one communication metaphor or channel in the human-computer interaction process could improve the usability of interfaces. Multimodal interaction could reduce the amount of information communicated by one specific metaphor or channel (Brown, Newsome et al. 1989) and increase the volume of communicated information (Oviatt 2003). Also, it enables users to interact with an application using the most suitable type of interaction metaphor to their abilities (Dix, Abowd et al. 2004). Multimodality also often makes the user interface more natural (Dix 1993) and offers conveying multiplicity of information through different channels (Sarter 2006).

The reviewed previous work demonstrated that using speech sounds and earcons improved the usability of computer applications in many different domains including e-learning. Also, it showed the usefulness of avatars with facial expressions in e-learning interfaces. However, there is a need for additional research to integrate multimodal metaphors in e-learning applications. Multimodal metaphors may help to alleviate some of the difficulties that e-learning users often encounter. This empirical study represents the initial experiment of a research project that aimed at exploring the usability aspects of e-learning interfaces that incorporate a combination of typical text with graphics metaphors and multimodal metaphors such as speech sounds, non-speech sounds and avatar with simple facial expressions. The main question is whether the inclusion of these metaphors can increase the usability and learning. The secondary question is related to the contributing role that each of these metaphors could play in enhancing the usability and learning of e-learning tools. The third question is whether the use of an avatar with human- like facial expressions and body gestures could help in improving the usability and learning. Finally, does it make difference between one avatar and two in the interface of e-learning applications? The purpose of this experiment was to answer the first question. An e-learning experimental platform, with two interface versions (a text with graphics and a multimodal), was developed to serve as a basis

for this study. The e-learning topic was class diagram notation that is often used in the design of software systems. The study involved two groups of users (one group for each interface version) in which the usability performance of the two groups in terms of efficiency, effectiveness, and user satisfaction was compared. The next two sections review the relevant literature in e-learning and multimodal interaction. Section 4 describes the experimental e-learning platform, and section 5 provides in detail the design of the experiment. The obtained results are presented and discussed in sections 6 and 7 respectively. Finally, section 8 concludes this chapter and indicates future work directions.

2. E-learning

In general, E-Learning could be defined as a term that describes the learning process in which information and communication technology could be utilized (Alexander 2001; Yu, Zhang et al. 2006) were computer machines and networks could be used to facilitate easier and faster access to a huge amount of educational content. Due to the continuous development of information and communication technology (ICT), research in e-learning and the technology employed in the development of e-learning applications have been increased (Hamilton, Richards et al. 2001). For example, Internet technology could be applied to manage the learning material and to store information about learners as well as to facilitate communication and cooperation between learners and to monitor their progress (Mikic and Anido 2006). Further technology used for e-learning include scheduled and on-demand delivery platforms (Hamilton, Richards et al. 2001). Scheduled delivery platforms were designed to simulate real learning environment but restricted by time and location constraints. Examples of this technology include video broadcasting, remote libraries and virtual classrooms. On the other hand, on-demand delivery platforms offer anytime and anywhere learning such as web based training and interactive training CD ROMs.

In comparison with traditional learning, e-learning offers better adaptation to individual needs and provides more flexible learning in terms of time and location, and ease monitoring student's knowledge and skills (Mikic and Anido 2006). Also, e-learning content could easily and quickly be updated and then redistributed again so that all users receive the same educational material in the same manner (Rosenberg 2001). Furthermore, it enables users to learn collaboratively (Correia and Dias 2001) and could enhance their motivation and interest in regard to the presented material (Theonas, Hobbs et al. 2008). Lastly, e-learning could be used to accommodate different teaching methods (Spalter, Simpson et al. 2000). However, e-learning environments also have challenges and difficulties. For example, users of e-learning are supervised only by parents or other adults not by a teacher. Also, teaching methods and computer technology must be combined appropriately (Russell and Holkner 2000) despite that this technology is not always available and accessible (Brady 2001). Furthermore, it was found that students, sometimes, are not satisfied with computer-based learning and experienced lack of traditional face-to-face contact with tutor. Hence, users' attitude towards e-learning should be enhanced and their accessibility to the needed technology should be facilitated (Shaw and Marlow 1999).

From pedagogical perspective, there are basic principles that must be considered to insure successful implementation of e-learning (Govindasamy 2001) . Govindasamy stated that development and evaluation of e-learning environments involves analysis of both learner and learning tasks, determination of instructional objectives and strategies, testing and

production of the initial version of the e-learning tool. It is not always the case that every e-learning system presents high quality learning. According to Henry (Henry 2001), a comprehensive e-learning solution consists of content, technology, and services. Also, the content of online courses as well as the tools used for producing this content should be carefully selected (Correia and Dias 2001). In addition to the ease of use, e-learning solutions should also be accessible without any hardware or software limitations (Kapp 2003). Moreover, the user interface of e-learning software should support independent learning, incorporate suitable communication metaphors, and provide a collection of pedagogical activities to suit the individual differences among users (Tabbers, Kester et al. 2005).
Human working memory is a short-term memory that plays an important role in the cognitive process for remembering and retaining the learning information that has been received (Graf, Lin et al. 2007). The capacity of this memory is limited and considered as one of the main cognitive features (Lin and Kinshuk 2003) therefore, when a large amount of information is presented via only one channel (i.e. visually), users will be cognitively overloaded (Low and Sweller 2005). It was found that the use of computers in learning could be advantageous over other media such as static and text-based material (Carswell and Benyon 1996). Also, involving other human senses than the visual one in e-learning interfaces will assist in extending the capacity of working memory and, as a result, users' ability to perceive and understand the presented information will be enhanced (Fletcher and Tobias 2005).

3. Multimodal Interaction

Multimodal interaction is a human-computer interaction in which more than one of human senses are involved through the incorporation of multimodal communication metaphors. These metaphors include text, graphics, speech sounds, non-speech sounds and avatars. Several studies have been carried out to examine how these metaphors could affect the interaction process. These studies showed that multimodal metaphors could be utilised to enhance the usability of many computer applications including e-learning.

3.1 Speech and Non-Speech Sounds

Sound is more flexible than visual output because it could be heard from all sides without paying visual attention to the output device. However, they are complement to each other and could be used simultaneously to transmit different types of information. Earcons are short musical sounds of non-speech nature (Blattner, Sumikawa et al. 1989) that has been used to involve the human auditory channel in the interaction with computer applications. It has been employed efficiently and effectively to improve the interaction with components frequently appeared in user interfaces such as scrollbar (Brewster, Wright et al. 1994) and progress bar (Crease and Brewster 1998). Also, it was successfully used to draw users' attention to events related to the development of program code (DiGiano, Baecker et al. 1993) and to communicate aspects of program execution (Rigas and Alty 1998). Other studies demonstrated that earcons could help in conveying auditory feedback to visually impaired users in order to communicate graphical information such as coordinate locations, simple geometrical shapes and their sizes (Alty and Rigas 2005; Rigas and Alty 2005). Moreover, it could be beneficial for sighted uses in accessing information represented in line graphs (Brown, Brewster et al. 2002), spreadsheets (Stockman 2004) and numerical data

tables (Kildal and Brewster 2005). In education, it was found that earcons could improve students' understanding (Bonebright, Nees et al. 2001) as well as their satisfaction in regards to the learning material (Upson 2002). Furthermore, earcons were successfully integrated with speech sounds to convey information to users (Rigas, Memery et al. 2001) and were found to have a positive contribution when included with recorded speech in the interface of a multimedia on-line learning tool as it helped users to perform different learning tasks more successfully (Rigas and Hopwood 2003).

3.2 Avatars

An avatar is another multimodal interaction metaphor that involves both of visual and auditory human senses. It is a computer-based character that could be utilised to represent human-like or cartoon-like persona (Sheth 2003) with the ability to express feelings, emotions and other linguistic information through facial expressions and body gestures (Beskow 1997). The six universally recognisable categories of human expressions are happiness, surprise, fear, sadness, anger and disgust (Fabri, Moore et al. 2002). It was found that adding an avatar with these facial expressions in the interface of Instant Messaging tools improved the involvement of users and created a more enjoyable experience for them (Fabri and Moore 2005). Also, a study conducted by Gazepidis and Rigas demonstrated the importance of specific facial expressions and body gestures when it has been used by a virtual salesman in an interactive system (Gazepidis and Rigas 2008). The role of avatar as a pedagogical agent in e-learning virtual environments has been evaluated by a series of empirical studies. Results of these studies suggested that avatar could facilitate the learning process (Baylor 2003; Holmes 2007), and could provide students with a sense of presence and enhance their satisfaction with the online courses (Annetta and Holmes 2006). What is more, the inclusion of avatar as a facially expressive virtual lecturer could lead to a more interesting and motivating virtual educational experience, and that the use of specific expressions (i.e. smiling) could increase students' interest about the learning topic and therefore enhance their learning performance (Theonas, Hobbs et al. 2008).

4. Experimental Platform

An e-learning platform was developed from scratch to serve as a basis for this empirical study. The platform provided two different versions; a text with graphics interface version, and a multimodal interface version. Both interfaces were designed to deliver the same information about class diagram representation of a given problem statement. The material presented by both interfaces, in the form of three common examples, included explanations about classes, associations among classes and the multiplicity of a given class in the diagram. The complexity of these examples was gradually increased, and each of which given in a separate screen display. Also, the order in which these examples were presented was similar in both interfaces..

4.1 Text with Graphics Interface

Figure 1A shows an example screenshot of the text with graphics interface. In this version of the experimental platform, the required information was delivered in a textual approach and could be communicated only by the visual channel without making use of any other

human senses in the interaction process. When the mouse cursor is placed over a given notation in the class diagram (denoted by 1), a textual description of that notation is displayed in the notes textbox (denoted by 2).

4.2 Multimodal Interface

Figure 1B shows an example screenshot of the multimodal e-learning interface. Guidelines for the design of multimodal metaphors (Sarter 2006) and multimodal user interface (Reeves, Martin et al. 2004) were followed. The same design of the text with graphics e-learning interface was used but the notes textbox was removed and replaced with a combination of recorded speech, earcons, and avatars with facial expressions. The life-like avatar with simple facial expressions (see figure 1C) was included to speak the explanations about classes with prosody. Multiplicity and notation of associations were communicated with earcons and recorded speech respectively. Two command buttons were also provided in both interfaces to allow users to select the three examples

Earcons employed in the multimodal interface were designed based on the suggested guidelines (Brewster, Wright et al. 1995; Rigas 1996). Musical notes (starting at middle C in the chromatic scale) were used to create six earcons each of which was utilized to communicate one of the six different types of multiplicity found in the three class diagram examples. Table 1 shows the design structure of these earcons. Each of the first four earcons were composed of two parts separated by a short pause (0.6 second) in between, and communicated one of the multiplicities: zero or 1 (0..1), one or more (1..*), two or more (2..*), and one or two (1..2). The remaining two earcons had only one part and used to represent the multiplicities: one (1) and many (*) which means zero or more. So, in order to create these six earcons, there was a need to musically illustrate the values 0, 1, 2, and *. For this purpose, different numbers of rising pitch piano notes were used as follows: one

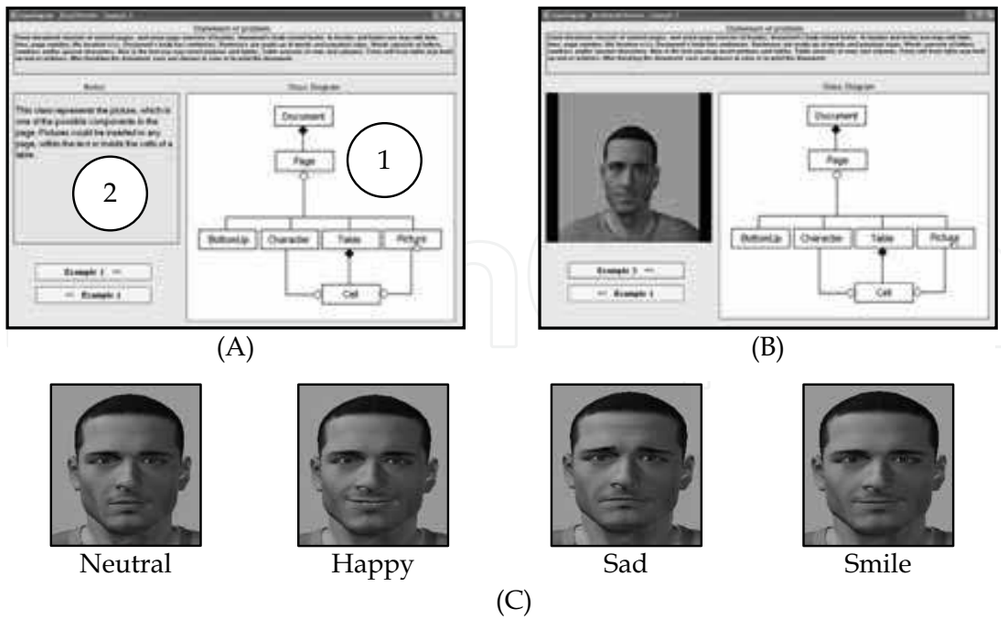


Fig. 1. A screenshot of text with graphics interface (A), multimodal interface (B) and the facial expressions used in the multimodal interface (C)

musical note to communicate 1, two rising notes to communicate 2 and four rising pitch notes to communicate many (*). In order to distinguish zero and to represent it in the multiplicity 0..1, only one note of seashore sound was used.

		Multiplicity					
		0..1	1..*	1..2	2..*	1	*
Duration (second)		1.8	2.2	1.9	2.4	1	1.2
Part 1	Sound	Seashore	Piano	Piano	Piano	Piano	Piano
	No. of tones	1	1	1	2	1	4
Part 2	Sound	Piano	Piano	Piano	Piano	-	-
	No. of tones	1	4	2	4	-	-

Table 1. The structure of earcons that communicated multiplicity in the multimodal interface

5. Design of The Two-Group Empirical Study

In order to explore the effect of multimodal metaphors and to find out which interface would be better in terms of efficiency, effectiveness, and user satisfaction for the e-learning process, the two e-learning interface versions were empirically evaluated by two independent groups each of which has fifteen users. One group used the text with graphics interface and served as a control and the other group used the multimodal interface in order to serve as experimental. Therefore, the main hypothesis stated that the multimodal e-learning interface would be more efficient, more effective and more satisfactory compared to a similar interface with only text with graphics metaphors.

Participants in the experiment were first-time users of the experimental platform. The majority of them in both groups were postgraduate students coming from a scientific background and had no or limited experience in class diagram notation. They were regarded as expert computer users because most of them use computer ten or more hours a week.

Both groups performed six common tasks. The tasks were designed to increase in difficulty and equally divided into easy, moderate and difficult. The tasks also covered all types of presented information such as class attributes and operations, associations between classes, and multiplicities. Each task comprised a set of requirements each of which asked the user to place the mouse cursor over a specific notation in the displayed class diagram, and to receive the delivered information related to that notation. The number of task's requirements depended on the complexity level of the task. Each task was evaluated with a memory recall and a recognition questions. To answer the recall question correctly, user had to retrieve part of the presented information from his/her memory. However, the recognition one offered a set of 2 to 4 options and user had to recognize the correct answer among it. In total, each user answered twelve questions consisted of 4 easy, 4 moderate and 4 difficult. In other words, these questions were categorised into 6 recall and 6 recognition questions. For the purpose of data collecting and analysis, each question was considered as a task. Table 2 shows the multimodal metaphors used to communicate the key information needed by users in the multimodal interface group to complete the tasks successfully.

	Tasks											
	Easy				Moderate				Difficult			
	Rl	Rn	Rl	Rn	Rl	Rn	Rl	Rn	Rl	Rn	Rl	Rn
	1.1	1.2	2.1	2.2	3.1	3.2	4.1	4.2	5.1	5.2	6.1	6.2
Speech		√		√		√		√				√
Earcons		√		√	√					√		√
Avatar	√		√				√		√		√	

Table 2. The multimodal metaphors used to communicate the key information needed by users in the multimodal interface group to complete the tasks successfully (Rl=Recall, Rn=Recognition)

The experiment was conducted individually for each user in both groups with an average duration of thirty five minutes. It started by filling the pre-experimental questionnaire for user profiling. Then, two tutorials were presented; the first tutorial demonstrated the class diagram notation for five minutes and was shown to each user in both groups. The second tutorial had two versions, one for each group. The aim of each of these tutorials was to provide an introduction to the e-learning interface version that the user was to use. Both of these tutorials run for two minutes. After completing all tasks, users were asked to give their satisfaction ratings about the different aspects of the tested interface version by answering the post-experimental questionnaire.

6. Results

The results of both groups were analysed in terms of efficiency (time users needed to accomplish the required tasks), effectiveness (percentage of correctly completed tasks) and user satisfaction (based on a rating scale). The mean completion time for all tasks (see figure 2B) in the experimental group was significantly lower than the control group ($t=1.74$, $cv=1.72$, $p<0.05$). Experimental observations revealed that users in the control group regularly divided their visual attention between the notes in the textbox and the class diagram representations in order to understand the presented information and in some cases a visual overload occurred. However, users in the experimental group maintained

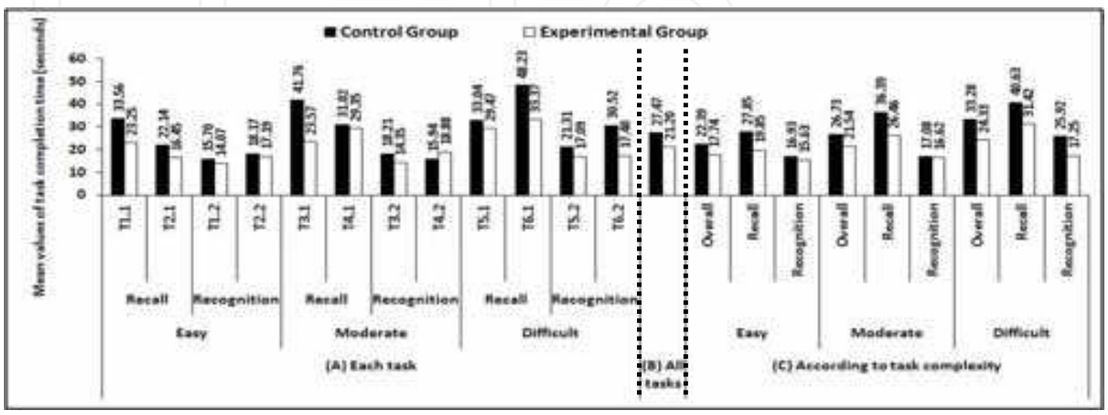


Fig. 2. Mean values of task completion time (A) for each task, (B) for all tasks, and (C) according to task complexity in both control and experimental groups

their visual attention to the class diagram representations while they were listening to the auditory messages.

A more detailed analysis of the mean completion time for each task in both groups is shown in figure 2A. The time users needed to complete the tasks in the experimental group was lower in 11 out of 12 tasks. It could be noticed that the difference between mean values of the two groups was varied over the twelve tasks. It was higher in four tasks (T1.1, T3.1, T6.1, and T6.2) which represent 33% of the tasks. However, it was lower in other four tasks (T2.1, T5.1, T3.2, and T5.2) and even more lower in the tasks T1.2, T2.2 and T4.1. Only task T4.2 recorded higher time when using multimodal interface without critically affecting the overall result. These variances in the difference between mean values of the two groups are attributed to the differences in the presented information and in the complexity of the required tasks. Furthermore, it was noticed that these difference did not clearly explained the role that each of speech, earcons and avatar played in enhancing the accomplishment time when used in the multimodal interface. The reason behind this could be returned to the design of the required tasks.

The tasks were designed to increase in difficulty and they were equally divided into easy, moderate and difficult. Figure 2C demonstrates that there is a relationship between the complexity of the tasks and the time required to complete the task. The completion time in the experimental group was lower for all tasks regardless the level of difficulty. However, the variance in completion time between the two groups increased as the task complexity increased. This demonstrates that the users in the experimental group were significantly aided by the multimodal metaphors. The t-test calculations were performed between the two groups to evaluate completion time for easy, moderate, difficult, recall and recognition tasks (see table 3). Values obtained in t-test calculations showed a statistical significant difference between the two groups in relation to the time users spent to complete recall tasks regardless of the tasks' complexity. The t values were 2.4 for easy recall, 2.25 for moderate recall, 1.93 for difficult recall and 3.94 for overall recall tasks. In recognition tasks, t values for easy (0.68) and moderate (0.17) tasks were not significant but significant was reached in the difficult tasks (3.40). Nevertheless, users in the experimental group spent significantly lower time than users in the control group to perform all recognition tasks.

	Easy			Moderate			Difficult			Rl	Rn
	Overall	Recall	Recognition	Overall	Recall	Recognition	Overall	Recall	Recognition	Overall	Overall
t value	**2.27**	**2.40**	0.68	**2.12**	**2.25**	0.17	**2.99**	**1.93**	**3.40**	**3.94**	**2.42**
Degree of freedom	118	58	58	118	58	58	118	58	58	178	178
Critical value	1.66	1.67	1.67	1.66	1.67	1.67	1.66	1.67	1.67	1.65	1.65

Table 3. Values obtained through t-test calculation when comparing the task completion time between the two groups. Statistically significant values are displayed in bold (Rl=Recall, Rn=Recognition)

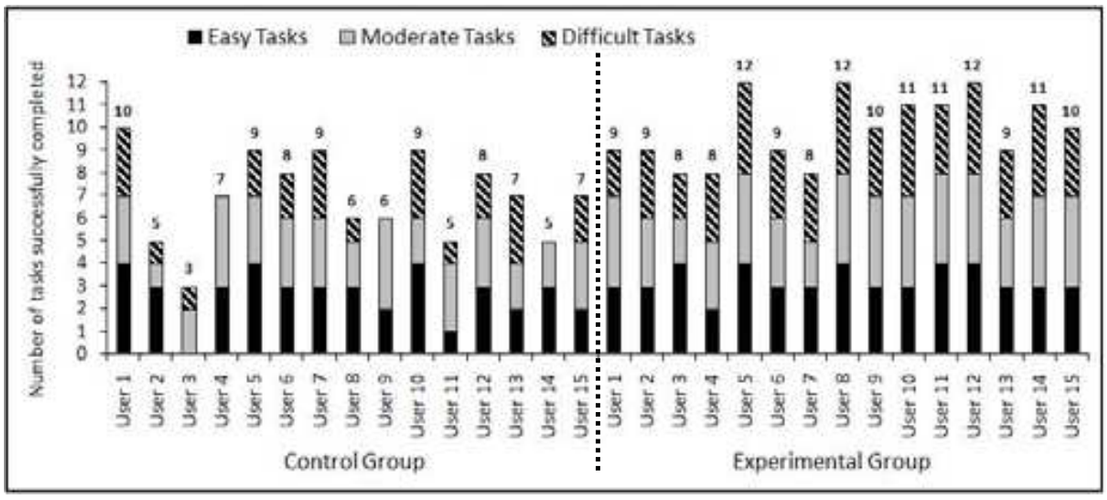


Fig. 3. The number of successfully completed tasks for each user in both groups: control and experimental

Figure 3 shows the number of correctly completed tasks (out of 12) for each user in both groups. The number of correctly completed tasks was higher in the experimental group as users successfully completed 149 (83%) out of 180 tasks. Users in the control group successfully completed 104 tasks (58%) out of 180 tasks. This difference was significant ($t=4.75$, $cv=1.70$, $p<0.05$).

Table 4 shows the mean values of successfully completed tasks for both groups and the t values obtained by t-test calculations (degree of freedom= 28, critical value= 1.70). The t values show that the successfully completed tasks in the experimental group are significantly higher regardless of the task complexity (easy, moderate, and difficult) or task type (recall and recognition). The only non significant difference was in the easy and moderate recognition tasks in which the multimodal metaphors used in the experimental interface did not contribute as much as in the other types of tasks.

			Easy			Moderate			Difficult			Rl	Rn
		All tasks	Overall	Recall	Recognition	Overall	Recall	Recognition	Overall	Recall	Recognition	Overall	Overall
Mean	C	6.9	2.7	1.1	1.5	2.7	1.2	1.5	1.6	0.7	0.9	3.0	3.9
	E	9.9	3.3	1.5	1.7	3.5	1.9	1.6	3.2	1.7	1.5	5.1	4.9
t value		**4.75**	**1.84**	**1.88**	0.81	**2.81**	**3.39**	0.57	**4.73**	**3.78**	**2.95**	**5.07**	**1.94**

Table 4. Mean values of successfully completed tasks in both control group (C) and experimental groups (E), and the t values obtained by t-test calculations (Rl=Recall, Rn=Recognition)

Figure 4 shows the percentage of tasks successfully completed in both groups. It can be noticed that the difference between the two groups in successfully completed tasks increased as the difficulty of the task increased (15% in easy tasks, 20% in moderate tasks

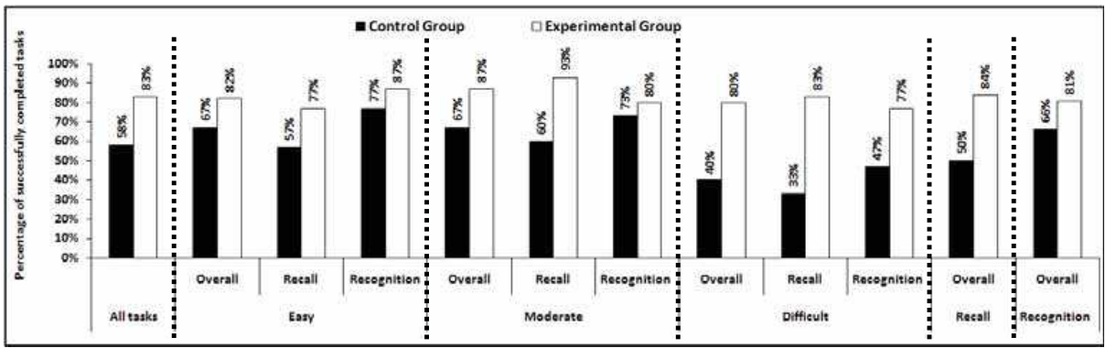


Fig. 4. Percentage of successfully completed tasks in control group and experimental group

and 40% in difficult tasks). Also, users of the multimodal interface performed better in both recall and recognition tasks but the difference between the two groups is smaller in the recognition tasks. Therefore, the contribution of multimodal metaphors, as used in the experimental interface, aided users to successfully complete more recall than recognition tasks.

User satisfaction in regards to different aspects of the applied interface was measured in both groups by users' responses to the post-experimental questionnaire. These aspects included ease of use, confusion, nervousness, ease of learning, identification and recognition of the learning material and overall satisfaction. A scale with six points was used for each statement in the questionnaire. This scale ranged from 1, the value of strongly disagree, to 6, the value of strongly agree. For the sake of data analysis, responses of each user were summated to obtain the satisfaction score for each user in each group. These total scores were then used in the statistical analysis and users in the experimental group were found to be significantly more satisfied ($t=2.76$, $cv=1.70$, $p<0.05$).

Figure 5A shows the overall mean values of user satisfaction score which was 4.4 in the control group and 5.2 in the experimental group. The multimodal e-learning interface was easier to use than the text with graphics one. It was observed that users in the experimental group were more relaxed, and less confused and nervous (see figure 5B).

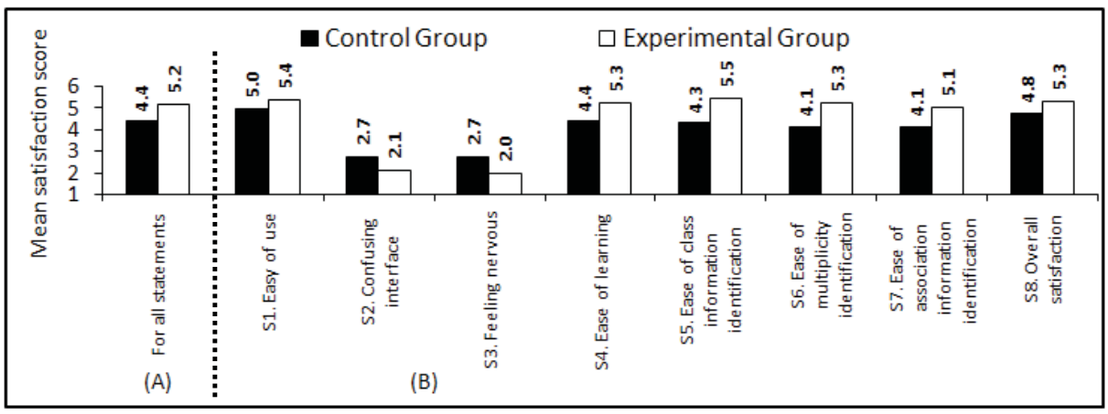


Fig. 5. Overall mean values of user satisfaction score for all statements (A) and for each statement (B) of the control and experimental groups

7. Discussion

The results of this two-group empirical study can be discussed in terms of the contribution of the multimodal metaphors to the usability and to the learning process of users. The analysis is presented from three angles:

1. Time taken to complete the various tasks in terms of type of tasks (recall and recognition) and in terms of the difficulty of the task.
2. Successful completion of tasks (again in terms of type and difficulty).
3. User satisfaction and experience in terms or ease of use, confusion of users, nervousness of users, ease of learning, identification and recognition of learning material, and overall satisfaction.

Although the text with graphics interface offered a simpler typical interaction, the results of the experiment showed that the use of multimodal metaphors (recorded speech, earcons, and avatars) was significantly more efficient and effective than using text with graphics to communicate information in an e-learning interface. Also, users who used the multimodal interface were significantly more satisfied than users who used the text with graphics interface.

7.1 Time Taken to Complete Tasks

During the experiment, it was noticed that users of the textual based interface switched their attention between the textual descriptions and the class diagram to understand the presented textual information, which may have overloaded their visual channel. On the other side, users of the multimodal interface were able to focus their attention on the class diagram while receiving information from the spoken messages and earcons. The inclusion of more than one communication metaphor in the multimodal interface helped users to concentrate better on the presented information through the auditory channel and at the same time use the visual channel to understand this information. The results of the study demonstrated that these metaphors assisted users to learn quicker especially as the required task became more difficult. The time spent on the completion of tasks increased as the task difficulty increased. However, the more important aspect is that the variance in completion time between the two groups increased as the task complexity increased. This demonstrates the contribution of speech, earcons, and avatar in users' efficiency of higher complexity tasks.

In recall tasks, users needed to retrieve the presented information from their memory and this may have taken time depending on the complexity of the task. Users in the experimental group used less time to complete easy, moderate and difficult recall tasks and difficult only recognition tasks. In other words, no significant difference between the two groups was observed for only easy and moderate recognition tasks. This means that the use of multi-modal metaphors as applied in the experimental interface particularly contributed to memory recall activities regardless of its complexity.

7.2 Successful Completion of Tasks

The use of more than one communication metaphor assisted users to distinguish among the different types of information provided by each of these metaphors and enabled them to remember this information for longer time. The fact that users in the multimodal group retained the communicated information for longer time (compared to the text with graphics group) enabled them to successfully complete more tasks

In order to successfully perform the recall tasks, users had to correctly retrieve from their memory the presented information. Information in the multimodal group was presented in a teacher like scenario in which the avatar simulated a teacher with natural head movement, facial expressions and natural speech while other aspects of the learning materials were presented using earcons. The results of the study demonstrated that the user experience as formed by combined multimodal metaphors enabled users to learn better. This is particularly demonstrated in the recall tasks which are more difficult to be completed than the recognition tasks (84% completion rate in the experimental group compared with 50% in the control group). The low completion rate of recall tasks in the text with graphics interface demonstrates that the users' memory was not aided as much as the multimodal interface. To perform recognition tasks successfully, users had to choose the correct answer among the given options. There is always a possibility that a correct answer could be chosen by the user due to chance (this is far more difficult to happen in a recall task). The successful completion was 66% in the text with graphics group and 81% in the multimodal group. The difference, although smaller than the one in the recall tasks, still indicates that users performed better when their e-leaning has taken place in the presence of multimodal metaphors.

In terms of task difficulty, the tasks were structured to gradually increase in complexity. Similarly to the results of time taken to complete the tasks, the results of the experimental study showed that the experimental group successfully completed more tasks as the level of difficulty increased. This is particularly noticeable in the difficult tasks were the difference was significant regardless of the tasks type (recall or recognition). In the moderate and the easy tasks, the difference was significant for recall and not for recognition.

7.3 User Satisfaction

The two interface versions (text with graphics and multimodal) did not demonstrate a significant difference for the ease of use, making users confused or nervous (see statements S1 to S3 in figure 5B). A larger difference however was observed on the specific statements relating to learning (see statements S4 to S7 in figure 5B). These results derived from two independent groups and users within those two groups were not presented with both interface versions in order to make an informed comparison. However, the users in the multimodal group may have had prior experience to typical learning interfaces and this probably served as a comparison point. Typically, users in the experimental group thought that their learning was better aided by the multimodal metaphors by one point more (in a 1 to 6 scale) than the users in the control group. Users easily identified learning information about classes, associations, and multiplicity, which communicated by avatar, speech, and earcons respectively. This result on its own is not conclusive as it is based on subjective rating of users and the typical mean difference is not large enough (although a statistical significance for the overall satisfaction results was reached). However, when the user satisfaction data is combined with the efficiency and effectiveness results, the argument that

users in the experimental group were helped by the multimodal metaphors becomes much stronger. It can therefore be extrapolated that multimodal aided learning, particularly in recall situations and complex learning material, is more likely to result in an enjoyable and satisfying experience for the user. This experience is linked with the ability to complete learning tasks correctly and quickly.

8. Conclusion and Future Work

In this chapter, we investigated the employment of multimodal metaphors (speech sound, non-speech sounds, and avatars with simple facial expression) as communication means to present information in the interface of e-learning platforms, and explored its effect on the usability of such interfaces. An experimental two-group study was conducted in which usability parameters: efficiency, effectiveness, and satisfaction of two different versions of the experimental e-learning platform were compared. In the first version, only visual channel of the users were used in the interaction to get textually presented information about class diagram notation. The second version of the experimental e-learning platform provided a combination of speech sounds (recorded), non speech sounds (earcons) and avatar with simple facial expressions to deliver the same information. The chapter then concluded with a discussion of the obtained results and research directions for future work.

The results obtained from this empirical study confirmed that the multimodal interface could indeed help to spend lower time in performing the required tasks and was more effective in conveying information in an e-learning platform. Also, it was more satisfactory. Nevertheless, these results did not clearly clarify the contributing role of each of earcon, speech and avatar. So, the inclusion of multimodal metaphors is suggested and should be taken into consideration when designing the user interfaces of e-learning applications. This will enable the user to make use of more than one channel of communication and, hence, shorten the time needed to perform the required tasks with higher level of accuracy. As a result, aspects of the usability in these interfaces will be improved.

Based on the experimental results and users' feedback, further enhancements on both of the experimental platform and the required tasks were needed to be used later in the next experimental work which will aim at investigating the contribution of each of speech, earcons, and avatar when used in the interface of an e-learning platform. In addition to the use of more than one avatar in the same e-learning interface, specific facial expressions and body gestures of a teacher-like avatar will be also explored.

9. References

Alexander, S. (2001). "E-learning developments and experiences." Education and Training **43**(4-5): 240-248.

Alty, J. L. and D. Rigas (2005). "Exploring the use of structured musical stimuli to communicate simple diagrams: the role of context." International Journal of Human-Computer Studies **62**(1): 21-40.

Annetta, L. A. and S. Holmes (2006). "Creating Presence and Community in a Synchronous Virtual Learning Environment Using Avatars." International journal of instructional technology and distance learning **3**: 27-43.

Baylor, A. L. (2003). The impact of three pedagogical agent roles. Autonomous Agents & Multi-Agent Systems Conference, Melbourne, Australia, ACM New York, NY, USA.

Beskow, J. (1997). "Animation of Talking Agents." Proceedings of AVSP **97**: 149-152.

Blattner, M. M., D. A. Sumikawa, et al. (1989). "Earcons and Icons: Their Structure and Common Design Principles." Human-Computer Interaction **4**(1): 11-44.

Bonebright, T. L., M. A. Nees, et al. (2001). "Testing the effectiveness of sonified graphs for education: A programmatic research project." International Conference on Auditory Display, Espoo, Finland.

Brady, L. (2001). "Fault lines in the terrain of distance education." Computers and Composition **18**(4): 347-358.

Brewster, S. A. (1997). "Using non-speech sound to overcome information overload." Displays **17**(3-4): 179-189.

Brewster, S. A. (1998). "The design of sonically-enhanced widgets." Interacting with Computers **11**(2): 211-235.

Brewster, S. A., P. C. Wright, et al. (1994). "The design and evaluation of an auditory-enhanced scrollbar." Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence: 173-179.

Brewster, S. A., P. C. Wright, et al. (1995). "Experimentally derived guidelines for the creation of earcons." Adjunct Proceedings of HCI **95**: 155-159.

Brown, L., S. Brewster, et al. (2002). "Browsing modes for exploring sonified line graphs." Proceedings of British HCI **2**: 2–5.

Brown, M. L., S. L. Newsome, et al. (1989). "An experiment into the use of auditory cues to reduce visual workload." Proceedings of the SIGCHI conference on Human factors in computing systems: Wings for the mind: 339-346.

Carswell, L. and D. Benyon (1996). "An adventure game approach to multimedia distance education." SIGCUE Outlook **24**(1-3): 122-124.

Correia, A. P. and P. Dias (2001). "Criteria for evaluating learning web sites: how does this impact the design of e-learning." Actas da II Conferência Internacional de Tecnologias da Informação e Comunicação na Educação: Desafios Challenges: 521-528.

Crease, M. and S. Brewster (1998). "Making Progress With Sounds-The Design and Evaluation Of An Audio Progress Bar." Proceedings of ICAD **98**.

DiGiano, C. J., R. M. Baecker, et al. (1993). "LogoMedia: a sound-enhanced programming environment for monitoring program behavior." Proceedings of the SIGCHI conference on Human factors in computing systems: 301-302.

Dix, A., G. Abowd, et al. (2004). Human-Computer Interaction (3rd Edition). Prentice Hall.

Dix, A. J. (1993). Human-computer interaction. New York; London, Prentice-Hall.

Fabri, M. and D. J. Moore (2005). "Is empathy the key? Effective communication via instant messaging." Proceedings of the 11th EATA International Conference on Networking Entities.

Fabri, M., D. J. Moore, et al. (2002). "Face Value: Towards Emotionally Expressive Avatars." In proceeding of AISB Convention on Animated Expressive Characters for Social Interaction, London, UK.

Fletcher, J. D. and S. Tobias (2005). "The multimedia principle." The Cambridge handbook of multimedia learning: 117–133.

Gazepidis, N. and D. Rigas (2008). "Evaluation of Facial Expressions and Body Gestures in Interactive Systems." International Journal of Computers, NAUN **2**(1): 92-97.

Govindasamy, T. (2001). "Successful implementation of e-Learning Pedagogical considerations." The Internet and Higher Education **4**(3-4): 287-299.

Graf, S., T. Lin, et al. (2007). "Analyzing the relationship between learning styles and cognitive traits." Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007): 235-239.

Hamilton, R., C. Richards, et al. (2001). An examination of e-learning and e-books.

Henry, P. (2001). "E-learning technology, content and services." Education+ Training **43**(4): 249-255.

Holmes, J. (2007). "Designing agents to support learning by explaining." Computers & Education **48**(4): 523-547.

Kapp, K. M. (2003). "E-learning basics: essay: The e-learning market: it's about the learner, not the instructor!" eLearn **2003**(6).

Kildal, J. and S. A. Brewster (2005). "Explore the Matrix: Browsing Numerical Data Tables Using Sound." Proc. ICAD2005.

Lin, T. and P. A. Kinshuk (2003). "Cognitive Trait Model-A Supplement to Performance Based Student Models." Proceedings of International Conference on Computers in Education, Hong Kong: 629-632.

Low, R. and J. Sweller (2005). "The Modality Principle in Multimedia Learning." The Cambridge handbook of multimedia learning: 147–158.

Mikic, F. and L. Anido (2006). "Towards a Standard for Mobile E-Learning." International Conference on Systems and International Conference on Mobile Communications and Learning Technologies, 2006. ICN/ICONS/MCL 2006. International Conference on Networking: 217-222.

Oakley, I., M. R. McGee, et al. (2000). "Putting the feel in 'look and feel '." Proceedings of the SIGCHI conference on Human factors in computing systems: 415-422.

Oviatt, S. (2003). "MULTIMODAL INTERFACES." The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications.

Reeves, L. M., J. C. Martin, et al. (2004). "Guidelines for multimodal user interface design." Communications of the ACM **47**(1): 57-59.

Rigas, D. and J. Alty (2005). "The rising pitch metaphor: An empirical study." International Journal of Human Computer Studies **62**(1): 1-20.

Rigas, D. and D. Hopwood (2003). "The Role of Multimedia in Interfaces for On-Line Learning." 9th Panhellenic Conference on Informatics (PCI'2003). , Thessaloniki, Greece.

Rigas, D., D. Memery, et al. (2001). "Experiments in using structured musical sound, synthesised speech and environmental stimuli to communicate information: is there a case for integration and synergy?" Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing: 465-468.

Rigas, D. I. (1996). Guidelines for Auditory Interface Design: An Empirical Investigation, PhD thesis, Loughborough University of Technology.

Rigas, D. I. and J. L. Alty (1998). "Using sound to communicate program execution." Proceedings of the 24th EUROMICRO Conference **2**: 625–632.

Rosenberg, M. J. (2001). E-learning-strategies for delivering knowledge in the digital age, McGraw-Hill.

Russell, G. and B. Holkner (2000). "Virtual schools." Futures **32**(9-10): 887-897.

Sarter, N. B. (2006). "Multimodal information presentation: Design guidance and research challenges." International Journal of Industrial Ergonomics **36**(5): 439-445.

Shaw, G. and N. Marlow (1999). "The role of student learning styles, gender, attitudes and perceptions on information and communication technology assisted learning." Computers & Education **33**(4): 223-234.

Sheth, R. (2003). "Avatar Technology: Giving a Face to the e-Learning Interface." The eLearning Developers' Journal.

Spalter, A. M., R. M. Simpson, et al. (2000). "Considering a full range of teaching techniques for use in interactive educational software: a practical guide and brainstorming session." Frontiers in Education Conference, 2000. FIE 2000. 30th Annual **2**.

Stockman, T. (2004). "The Design and Evaluation of Auditory Access to Spreadsheets." Proceedings of ICAD.

Tabbers, H., L. Kester, et al. (2005). "Interface design for digital courses." In W. Jochems, J. van Merriënboer, & R. Koper (Eds). Integrated e-learning: implications for pedagogy, technology & organisation (pp. 100-111). London: Routledge Farmer.

Theonas, G., D. Hobbs, et al. (2008). "Employing Virtual Lecturers' Facial Expressions in Virtual Educational Environments." International Journal of Virtual Reality **7**(1): 31-44.

Upson, R. (2002). "Educational sonification exercises: Pathways for mathematics and musical achievement." International Conference on Auditory Display, Kyoto, Japan.

Yu, D., W. Zhang, et al. (2006). "New Generation of E-Learning Technologies." Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06).



Human-Computer Interaction
Edited by Inaki Maurtua

ISBN 978-953-307-022-3
Hard cover, 560 pages
Publisher InTech
Published online 01, December, 2009
Published in print edition December, 2009

In this book the reader will find a collection of 31 papers presenting different facets of Human Computer Interaction, the result of research projects and experiments as well as new approaches to design user interfaces. The book is organized according to the following main topics in a sequential order: new interaction paradigms, multimodality, usability studies on several interaction mechanisms, human factors, universal design and development methodologies and tools.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Marwan Alseid and Dimitrios Rigas (2009). An Empirical Investigation into the Use of Multimodal E-Learning Interfaces, Human-Computer Interaction, Inaki Maurtua (Ed.), ISBN: 978-953-307-022-3, InTech, Available from: http://www.intechopen.com/books/human-computer-interaction/an-empirical-investigation-into-the-use-of-multimodal-e-learning-interfaces

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2009 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen