Behaviour & Information Technology
Vol. 27, No. 5, September–October 2008, 395–405

Taylor & Francis
Taylor & Francis Group

Do video clips add more value than audio clips? Presenting industrial research and development results using multimedia

Bruce Christie and Jenny Collyer*

Department of Computing, Communications Technology and Mathematics, London Metropolitan University, London City campus, 31 Jewry Street, London, EC3N 2EY, UK

(*Received December 2004; final version received July 2006*)

It was hypothesised that in multimedia information applications a visual component can add value to otherwise audio-only clips. Subjects rated clips used in published CD-ROMs and how much they remembered was also tested. In one set of clips, the visual component was removed, in a second the audio was removed and a third set was unedited. The experiment was run three times, on different groups, to check replicability. For all groups, clips with a visual component as well as audio were judged to contain more information and to be more interesting than audio-only clips. There was also some evidence that the visual component can increase subjects' confidence in what they can remember. Other expected effects were not observed in every group in which they were tested. It is speculated that training in media literacy may help developers to use the visual component more effectively.

Keywords: audio; video; multimedia; presentations

1. Introduction

The applied context for the present study comes from the emergence of multimedia. Multimedia technology is continuing to expand the possibilities for communication and presentation of information in practical situations. Potential applications in business have been identified for some time (e.g. Bull *et al*. 1995). Some companies now issue multimedia CD-ROMs along with conventional marketing literature. Many now routinely include their web site addresses in their printed advertisements (see Pardun and Lamb 1999) and the web sites referred to increasingly make use of multimedia (see Hashmi and Guvenli 2001). Multimedia is also being applied in distance learning (e.g. Tiernan and Grudin 2000), as well as in a variety of other applications, including, for example, interactive maintenance manuals (e.g. Brinkman *et al*. 2001), the visualisation of financial information (e.g. Records and Olinsky 1998), meetings (e.g. Panteli and Dawson 2001) and many other applications in business, in education, at home and in other contexts.

Within this general context, the present study focuses in particular on the use of video in multimedia. Whether using CD-ROM, DVD, the Internet or other means of delivery, multimedia will often include video. Video is appealing. The success of television advertising owes much to this. Video seems intuitively to offer, as Meisel (1998) puts it: 'immediacy and a singular ability to focus attention on its subject'. Perhaps, as he suggests, this is because our cultural norms lead us to pay attention to video. There are, in any case, cognitive advantages in including a visual component in communication, supporting the principle (Mayer 2001) that combining words and pictures can be more effective in facilitating learning than using only words. It is known, for example, that the number of cues and channels used to receive information, along with other factors, can be important in reducing uncertainty and ambiguity (e.g. Panteli and Dawson 2001). Consistent with this, adding visual images to a presentation can be helpful in these respects. For example, in the context of a knowledge-based, multimedia system designed to train users to prepare cost estimates for building projects, Shen *et al*. (2001) feel that the system benefits users by being able to provide visual images (e.g. of different types of wall), along with additional explanations in the form of voice and video clips. They refer to the old adage that a picture is worth a thousand words.

Is a picture always worth a thousand words? There is some evidence to support this view. Even still pictures can often enhance the effectiveness of text (e.g. Carney and Levin 2002, Schnotz 2002). Animation added to a narration can sometimes enhance learning (e.g. Mayer 2001), although not always in every respect (e.g. Mayer 2001). Video can also sometimes be helpful. For example, Furnham *et al*. (2002) found that their 11 – 23 year-old subjects could recall more from news stories presented in their original televised

*Corresponding author. Email: j.collyer@londonmet.ac.uk

ISSN 0144-929X print/ISSN 1362-3001 online

DOI: 10.1080/01449290600959146
http://www.informaworld.com

form than from a print version. However, it is not clear from that experiment how important the visual component of the television version was compared with the audio component. A study by Faraday and Sutcliffe (1997) suggests that the visual component can be important. They showed that a multimedia presentation could be more effective in promoting learning than the same presentation with all the multimedia elements removed except for text and speech. In their study, however, the multimedia had been specifically redesigned in order to facilitate learning, taking account of eye movement patterns the experimenters had observed. Also, it is not known to what extent, in their combined text and speech condition, the speech was important compared with just the text. Other evidence (e.g. Matarazzo and Sellen 2000) suggests that, instead of being of benefit, video can sometimes have a distracting effect. This is consistent with the view (e.g. Nagy *et al*. 1999) that different media can sometimes interact with each other to produce emergent properties that can be difficult to predict.

Even where a net benefit of adding video can be expected, it is necessary in practical applications to take account of the technical limitations of the technology. In particular, at any given stage of its development, limitations of the technology may affect different cues differentially. It may be, for example, that, in the case of motion video, the loss of temporal information that results from the limited frame rate often associated with today's multimedia might have differential effects on different aspects of communication. It might not be too much of a problem in terms of some aspects of communication, such as recognising who is speaking, for example, but it may be a problem in relation to interpreting facial expressions, eye gaze and in supporting speech perception (Bruce 1996). In addition, as Panteli and Dawson (2001) point out, it is conceivable that, in practice, contextual factors may intervene to restrict or outweigh possible effects of information richness. In their study of meetings held using a visual conferencing system, there was evidence that their subjects would have needed to have learned some new behaviours appropriate to the medium in order to maximise its communication potential. This echoes comments made by others in relation to other applications. For example, in the case of financial imaging systems using multimedia, Records and Olinsky (1998) suggest that managers may need to change their expectations about information and how it is presented if they are to make effective use of financial imaging. These examples underscore how intuitive qualities of different media, even if theoretically plausible, need to be put to empirical test.

The benefit of video may therefore not be so dramatic or so unambiguous as it might seem intuitively. This is important because, in practical applications, the value, if any, of adding video to the application has to be balanced against its cost compared with other media. In particular, the cost of producing clips containing both visual and audio components is usually much higher than the cost of audio alone. This stems from a number of factors, including payment of fees to the directors, artists and technicians involved, intellectual property rights, production times, cost of the equipment involved and other factors. Where the video is made available over a network, such as the Internet, there is also the cost in terms of bandwidth to consider. Many Internet users currently use dial-up or other communication links that restrict the amount of information that can be transmitted in unit time, severely restricting the use of video. In their consideration of the prospects for multimedia content on the web, Hashmi and Guvenli (2001) suggest that the biggest advantage of audio in this context (compared with video) is its accessibility by a majority of users. Given the additional costs of video, it is reasonable to ask what value, if any, the visual component actually adds in practice.

The study reported here was specifically concerned with one aspect of using video in multimedia: the value of adding a visual component to what would otherwise be audio-only clips. It was concerned with the value that a visual component might add in practice, in commercially produced multimedia, rather than in experimental material. In contrast to the Faraday and Sutcliffe study (1997), referred to above, the present study used commercially produced multimedia material designed for marketing purposes and not redesigned by the researchers. The multimedia concerned had been produced for use in marketing the results of industrial research and development. In that context, including video in the multimedia might have value in several ways. For example, it might have value in terms of purchasers and users of the product feeling that they are being presented with more information (whether or not they can remember it). It might also make the information seem more interesting. It might hold their attention more and encourage them to revisit it. This interest in the more general impact on purchasers and users is consistent with a movement towards greater user-orientation in the communication of information, even technical information (cf. Zachry *et al*. 2001).

The aim of the study was to test the hypothesis that adding a visual component to what would otherwise be purely audio items can add value to them. Seven predictions consistent with the hypothesis were formulated, covering three main areas: effects on observers' interest; effects on the perceived amount of

information being presented; and effects on memory. These three types of impact on the observer were selected for their intrinsic interest given that the clips had been developed in order to be of value to those interested in research and development in information and communication technologies (ICT). The clips could reasonably be considered relatively unsuccessful if they failed to interest those with an interest in ICT, were perceived to convey little information or were such that the information presented was difficult to remember. A total of seven predictions were made. It was predicted that adding a visual component to what would otherwise be purely audio items would:

(1) result in the subjects giving higher estimates of the number of units of information contained in the items concerned;
(2) increase the subjects' estimates of the proportion of information in the items that could be remembered;
(3) reduce the subjects' estimates of the proportion of time their attention wandered;
(4) increase the subjects' ratings of how interesting the items were;
(5) increase the subjects' ratings of their interest in having the items presented again;
(6) improve the subjects' scores on an objective test of memory for information contained in the audio component;
(7) increase the subjects' ratings of their confidence in their answers to the objective test of memory.

Together, these seven predictions address both objective and subjective aspects of the possible value of adding a visual component to what would otherwise be audio-only material. Both objective and subjective aspects are important in the design of multimedia. It seems reasonable to suppose that, even if the material has value in objective terms, it may not be used if it fails to deliver in terms of users' subjective experience. Conversely, users may find some (subjective) value in the material if it seems interesting to them and appears to contain a lot of information that they feel confident they can remember afterwards even if, objectively, it does not.

2. Method

2.1. The clips used

The clips were taken from published interactive multimedia CD-ROMs presenting the results of a wide range of research and development projects in information technology. The CD-ROMs were published as part of a programme aimed at disseminating the project results to interested parties, especially technology brokers and other intermediaries playing key roles in the commercial exploitation of research and development. All the clips had been made by multimedia development companies on a commercial basis. The clips were used in preference to other sets that might, in principle, have been available because: (a) they had been made in a commercial context and had a degree of ecological validity that experimental materials typically lack; (b) their subject matter was consistent with the interests of the students who were available for the research; (c) they were readily available to the researchers.

Two sets of clips were used. Set 1 was 16 clips selected from a collection of 38. Set 2 was 19 clips selected from a second collection in the same series, containing 35 clips. The clips were selected at random with the constraint that none should be shorter than 55 seconds or longer than 85 seconds. (The clips in the original collections ranged from 44 seconds to 116 seconds in length, except for one exceptionally long clip that was 175 seconds.)

In set 1, five clips were selected at random to be presented with both the audio and visual components included (the 'audio plus video' condition; AV). Five others were selected at random to be presented with the visual component absent (the 'audio-only' condition; A), and another five were shown with the audio component absent (the 'video-only' condition; V). A single, quasi-random order of the 15 clips was used for the two groups (see below) to whom the set 1 clips were shown. The order was random, with the constraint that the same condition should not occur more than three times consecutively. The remaining clip (selected at random from the set of 16) was referred to as 'the reference item' and was shown before the first of the 15, before the sixth and before the 11th, thereby dividing the set of 15 into three blocks of five. The reference item was shown with both the audio and visual components included.

In set 2, one clip was selected at random to be the reference item. The other 18 were arranged in six blocks of three, each block containing the AV, A and V conditions. The order of the conditions for any given block was random, with the constraint that all six possible orders were used in creating the six blocks. The reference item was shown at the start and then again before blocks 3 and 5.

2.2. The forms the subjects completed

The subjects were each asked to complete two one-page forms immediately following each item except the reference item. The first form asked five questions requiring the subject to make a judgement, as follows:

(1) *Enter a number to complete the following sentence:*
Given that the reference item in the form it was presented by definition contains 1000 (a thousand) units of information, then I estimate that the item just presented in the form it was presented contains _____ units of information.

(2) *Enter a number between 0 (not at all) to 100 (all of it):*
I estimate that I can remember _____% of the information in the item.

(3) *Enter a number between 0 (none of the time) to 100 (all of the time):*
I estimate that my attention wandered from the item _____% of the time.

(4) *Enter a number (0 to 10) to complete the following sentence:*
I give the item just presented _____ marks out of 10 for how interesting I found it.

(5) *After all the items have been presented, one will be presented again, according to the wishes of the audience as indicated here. Enter a number from 0 (not at all) to 10 (maximum) to indicate how much you would like the item just presented to be presented again at the end:*
I give the item _____ out of 10.

The second form made two statements about the research and development result concerned, based on information presented in the clip. The statements were all based on the transcript of what was said in the audio, but this was not made explicit to the subjects. The subjects were asked to consider whether: 'Based on the item just presented, are the following statements true or false?' and to tick a 'true' or 'false' box for each question accordingly. They were instructed: 'If you are not sure whether a statement is true or false, guess'. An example of the statements made is as follows:

'The key benefits are that reliability is improved and costs are reduced without affecting the range of services or the level of security, which remain unchanged'.

The statements were based on parts of the transcript selected at random. Each statement was written by one of the experimenters. Its truth or falsity was checked by the other experimenter completing the form after comparing the statement with a printout of the relevant parts of the transcript. The second experimenter was not told whether the statement had been written with the intention of it being true or false. In a very few cases, the second experimenter's judgement of the truth or falsity of the statement differed from the intended answer. In those few cases the statements were re-written and tested on other colleagues until their truth or falsity was clearly established.

The form for group 3 also included a 7-point scale (from 'not confident at all' to 'extremely confident') for the subjects to record, separately for each answer, their confidence that their answers were correct.

2.3. Subjects

A total of 42 students studying information technology participated in the research as part of their studies. Group 1 were 11 postgraduates studying for an MSc in Multimedia Systems, group 2 were 18 undergraduates studying for a BSc in Multimedia Systems and group 3 were a mixture of 13 postgraduates studying for an MSc Multimedia Systems or an MSc User-Interface Design. The three groups were selected because they all had an interest in the type of information in the clips used in the study (about information technology). They participated as three separate groups, rather than one, for logistical reasons associated with their courses of study. All three groups were included, rather than just one of them, in order to assess the robustness of the findings across different groups (with slightly different experimental procedures, as explained below).

2.4. Procedure

Each group of subjects viewed the clips projected from a computer onto a large screen in a lecture room. Groups 1 and 2 were presented with set 1. For those groups, the various conditions were created by switching the audio or video on or off at the computer and projection system. Group 3 were presented with set 2. For that group, the different conditions were created by editing the clips and embedding them in a computer presentation package. This was done to reduce the load on the experimenter during the presentation of the clips. Instead of having to refer to a list showing the order of presentation and then switch the audio or video on or off accordingly, the order was programmed into the presentation software and all the experimenter had to do was go to the next clip.

At the start, the subjects were told they would be presented with a series of clips taken from an interactive multimedia application concerned with the results of research and development in information technology. The clips would be a mixture, some having both audio and video, some just audio and some just video. They were not told the hypothesis that was to be tested. They were told they would be asked to complete two simple forms after each item was presented. They were given a copy of the first form to familiarise

themselves with before the start, and the nature of the second form (the true/false test) was explained. They were not told what predictions were being made. They were then shown the reference item and it was explained that the reference item would be shown a couple of times more during the sequence, to refresh their memories of it.

Following each clip except the reference item, each subject completed the two forms and handed them in before the next clip was shown. The forms were completed anonymously but each subject was given a letter that he/she wrote on the forms. It was explained that this procedure would enable the subjects to respond anonymously whilst allowing the researchers to keep related forms together.

Groups 2 and 3, having had all the items presented to them, were asked to complete the second form (the true/false test) for two clips they had not been presented with. It was explained they would need to do their best and guess the answers if need be.

3. Results

The results comparing items having a visual component as well as an audio component with items having only an audio component are summarised in Table 1. The table is divided into seven sections, numbered 1 to 7, in order of the seven predictions made. Each section presents the results for a particular prediction. The remaining tables present supplementary results that are considered in the discussion.

Scores on the test of memory compared with guessing are summarised in Table 2.

The product-moment correlations between objective performance on the test of memory and the proportion of information the subjects felt they could remember are summarised in Table 3.

The product-moment correlations between objective performance on the test of memory and the subjects' confidence in their answers are summarised in Table 4.

Table 1. Results comparing items having a visual component as well as an audio component with items having only an audio component. (No. Ss is Number of Subjects.)

		A visual component					
		Present		Absent			
Group	No. Ss	Mean	SD	Mean	SD	t	p (one-tailed)
1 – Estimated amount of information in the item							
1	11	583	511	423	320	2.25	0.024*
2	18	882	677	518	402	2.94	0.005*
3	13	888	609	364	150	3.18	0.004*
2 – Proportion of information subjects felt they could remember							
1	11	49	19	43	19	1.69	0.061
2	18	56	13	45	15	3.69	0.001*
3	13	47	27	46	27	0.08	0.467
3 – Proportion of time attention wandered							
1	11	17	14	24	17	−2.25	0.024*
2	18	23	11	32	17	−2.69	0.008*
3	13	19	17	17	11	0.64	0.268
4 – How interesting the items were judged to be							
1	11	5.0	1.2	4.0	1.4	3.13	0.005*
2	18	5.8	1.6	4.7	2.0	3.23	0.002*
3	13	4.4	2.5	3.2	2.0	2.70	0.010*
5 – Interest in seeing or hearing the item again							
1	11	2.7	1.8	2.3	1.8	1.13	0.142
2	18	3.8	2.0	3.5	2.1	0.66	0.258
3	13	3.9	2.5	3.1	2.4	1.91	0.040*
6 – Number of items correct on the objective test of memory							
1	11	8.8	1.2	6.7	1.2	4.08	0.001*
2	18	6.8	1.4	6.3	1.3	1.34	0.099
3	13	9.4	1.6	9.9	0.8	−1.10	0.145
7 – Confidence in the answers given on the test (data collected only for group 3)							
3	13	4.9	1.1	4.3	1.3	2.77	0.008*

*The difference is significant at the 0.05 level.

Table 2. Scores on the objective test of memory for items with only a visual or only an audio component compared with guessing (average scores per item) (data collected only for groups 2 and 3).

			Conditions			Comparisons: p values (one-tailed) and t values	
Group	No. Ss		Guess G	Visual only V	Audio only A	V-G	A-G
2	18	Mean	0.86	1.14	1.27	$p = 0.011$*	$p = 0.000$*
		SD	0.38	0.28	0.26	t = 2.53	t = 4.10
3	13	Mean	0.96	1.04	1.64	$p = 0.277$	$P = 0.000$*
		SD	0.32	0.31	0.13	t = 0.61	t = 7.00

*The difference is significant at the 0.05 level.

Table 3. Product-moment correlation between the amount of information subjects felt they could remember and their scores on the objective test of memory.

			Components present		
Group	No. Ss		Both audio and visual	Audio only	Visual only
1	11	Correlation, r	0.36	0.23	0.11
		One-tailed p	0.140	0.248	0.372
2	18	Correlation, r	−0.06	0.17	−0.07
		One-tailed p	0.404	0.249	0.394
3	13	Correlation, r	−0.24	0.11	−0.20
		One-tailed p	0.215	0.357	0.260

Table 4. Product-moment correlation between the subjects' scores on the objective test of memory and their confidence in their answers (data collected only for group 3).

			Components present			
Group	No. Ss		Both audio and visual	Audio only	Visual only	None (Guess)
3	13	Correlation, r	0.48	0.70*	0.20	0.32
		One-tailed p	0.050	0.004	0.255	0.145

*The correlation is significant at the 0.05 level.

4. Discussion

The results taken as a whole (specific results are considered below) support the hypothesis that adding a visual component to what would otherwise be purely audio items can add value to them in terms of their impact on observers. However, the value may be more limited and less easy to demonstrate than might be expected in that the results were in line with the predictions to varying extents, as follows.

The results were fully consistent with three of the seven predictions, in that they were as predicted for all three groups or (in one case) for the only group in which the prediction was tested. Adding a visual component to what would otherwise be purely audio items:

- resulted in higher estimates of the number of units of information contained in the items concerned (prediction 1);
- increased how interesting the subjects consider the items to be (prediction 4);
- increased the subjects' confidence in their answers to the objective test of memory (prediction 7) in the only group where that was tested.

The results were also consistent with all of the other predictions for at least one of the groups on which they were tested. However, they were not consistent with the predictions for all the groups and so, in this sense, they are not as clear-cut as those above. Adding a visual component to what would otherwise be purely audio items:

- reduced the subjects' estimates of the proportion of time their attention wandered (prediction 3) – two of the three groups;
- increased estimates of the proportion of information that could be remembered (prediction 2) – one group;

- increased interest in having the items presented again (prediction 5) – one group;
- improved scores on an objective test of memory for information contained in the audio component (prediction 6) – one group.

The findings and their implications are discussed in what follows.

4.1. Subjects' judgements of the items (predictions 1 and 4)

The results clearly support the hypothesis that adding a visual component to what would otherwise be audio-only items adds something in terms of subjects' judgements of the items concerned. In particular, the items with a visual component were judged to contain more information (prediction 1) and to be more interesting (prediction 4). These two effects were statistically significant for all three groups. They provide the main justification in terms of the evidence from this study for including audio-video clips in multimedia productions rather than less expensive audio-only clips.

4.2. Subjects' ability to remember information presented (prediction 6)

The clips concerned were intended for use primarily by people seeking information to use in a professional or business context, as well as being of value to others, including researchers and students. This suggests that it would be appropriate for the clips to have been designed to help users to remember the information in them. In particular, it would be helpful if the visual component added value in assisting with remembering what was in the clips. It is therefore interesting that there was only modest evidence for such an effect in the present study.

Group 1 scored significantly better in the objective test of memory for those items that had a visual as well as an audio component (see Table 1, prediction 6). This suggests that the visual component may have helped that group. However, the effect was not replicated in the other two groups. Indeed, for group 3, the difference between the means was in the opposite direction to that predicted, but the difference did not approach statistical significance. It is possible that the statistically significant result for group 1 was a statistical anomaly. (A difference between means that meets the 0.05 threshold for statistical significance can be expected from random numbers about five times in 100, on average.) Alternatively, it is possible that adding a visual component does facilitate memory but that its effect is relatively small and/or relatively unpredictable, possibly being affected by a number of interacting factors (not necessarily very obvious factors) that may have varied from group to group in the present study. Either way, it seems reasonable to say that, in the present study, it did not prove possible to easily demonstrate a large effect of the visual component on memory that could readily be repeated using different groups of subjects. This is not to say that, with large samples and a powerful statistical test, it would not be possible to demonstrate an effect repeatedly – simply that any effect there was in the present study was not obvious enough to show up in all three groups.

The fact that an effect of adding a visual component was not more clearly demonstrated in all three groups is especially interesting as it is known (Eysenck 1998) that subjects can sometimes improve their scores on memory tests when they know what kind of questions they will be asked when they are being exposed to the material they have to remember. In the present study, especially after the first few clips, the subjects could be expected to have developed a good idea of the kinds of questions they would be asked. One might speculate, therefore, that the value of adding a visual component might be further reduced in situations in which the subjects are not aware of what they might be called upon to remember. Further research would be needed to test this possibility in the context of multimedia.

In principle, one possible explanation for why it did not prove much easier to demonstrate an effect of the visual component on memory for all three groups is that the visual component is inherently limited in what it can do to help users to remember the information in the audio. An alternative possibility is that, whilst adding a visual component can have a positive effect, the extent to which it does in practice depends, in part, on how it is used. This point is made by Nagy *et al.* (1999) and Bobrowicz and Christie (2003) in their discussion of media literacy and the emergent properties of multimedia. One might speculate that for the visual components of the clips used in the present study to have had a more obvious effect on memory for information in the audio components, they would have had to have been used more effectively, perhaps reflecting greater media literacy. This is also consistent with Mayer's (2001) point concerning the effect of design on cognitive load. Following Sweller and Chandler (1994) and Sweller (1999), he argues for a distinction between intrinsic cognitive load (determined by the inherent difficulty of the material being presented) and extrinsic cognitive load (determined by the way the message is designed). He suggests that whilst good design can minimise the extrinsic cognitive load on the learner, poor design can have the opposite

effect, especially by requiring the learner to engage in irrelevant or inefficient cognitive processing. Given that the clips used in the present study were produced in a real-world commercial context, the present results could be taken as an interesting example of a possible need for multimedia designers to receive more or better training in media literacy, especially with regard to the impact of design on extrinsic cognitive load. This point is taken up again later in this paper.

4.3. Subjects' feelings about how much they could remember (prediction 2)

Consistent with the evidence above that adding a visual component helped subjects, at least in group 1, to remember the information when their memory was tested objectively, there is some evidence that subjects also felt they could remember more when there was a visual component as well as audio. This can be seen in table 1 (prediction 2), which shows that group 2's estimates of how much they could remember were significantly higher for clips that had a visual component as well as audio. Their performance in the objective test did not reflect this. (See Table 1, prediction 6, where the difference between the conditions for this group was no greater than would be expected by chance.) So either they were mistaken or the test was failing to tap the kind of additional information they were thinking of. The latter is quite likely as they were asked to estimate what proportion of 'information in the item' they could remember, not how much in the audio component specifically (which is what the objective test was based on). Whatever the exact nature of the effect, it was not replicated in the other two groups.

Table 3 shows there was no significant relationship between how much subjects felt they could remember and what they actually scored on the memory test. This is true for all three groups and for all types of clip. If, in Table 3, only the results for the clips with both visual and audio components are considered, they might suggest that how much subjects could remember from the visual component is unrelated to how much they could remember from the audio component. (Their estimates might reflect an assessment of how much they felt they could remember from both, whereas the objective test only measured how much they could remember from the audio.) However, this would suggest that their estimates would be consistent with the results of the objective test for clips with only one component. Table 3 shows this was not the case. It seems from the present study that subjects are simply not very good at estimating how much information they can remember from such clips. In terms of assessing the value of adding a visual component, this means that clips that have value in terms of the subjects' subjective feeling about how much they can remember may or may not have equivalent value in objective terms.

4.4. Subjects' confidence in the accuracy of what they could remember (prediction 7)

The subjects in group 3 were asked to rate their confidence in their answers in the memory test. Table 1 (prediction 7) shows that they reported feeling significantly more confident when asked about items with a visual component as well as an audio component than when asked about audio-only items. However, Table 1 (prediction 6) also shows that, for this group, adding a visual component did not improve scores on the memory test any more than might be expected by chance when their memory was tested objectively. In fact, the (statistically non-significant) difference between the two conditions was in the opposite direction. It seems that adding a visual component may lead subjects to be more confident in their answers than an objective test of memory suggests they should be. In that respect, adding a visual component could be considered to at least carry a risk of having a negative value. However, as confidence was measured only in one group, further research is needed to assess whether the finding can be repeated across a range of other groups and conditions.

Table 4 shows that subjects who scored better in the memory test also tended to be more confident in their answers when the clips contained the audio component (with or without the visual component). The table also shows that the relationship did not apply when the audio component was absent (the 'visual only' condition and the 'guessing' condition). It would seem that, whatever information subjects were responding to in the visual component that influenced their feelings of confidence, it did not actually help them to answer the questions they were asked. The lack of a significant relationship between confidence and test score for the guessing condition supports the view that the questions in the memory test did not themselves contain useful information about what the answers were. (As noted above, there was a positive relationship between confidence and test scores when relevant information was available – in the audio component. Given that, then one would expect that, had the questions hinted at the correct answers, subjects who were better able to pick up on those hints would have tended to have scored better and been more confident than those who did not.)

4.5. Subjects' feelings about their own behaviour (predictions 3 and 5)

In addition to the effects noted above, the findings suggest there may be two other effects on subjects' feelings. First, in two of the groups, subjects felt that their attention wandered less from the items with a visual component (prediction 3). Second, in one of the three groups, they expressed more interest in having the items with a visual component presented again, rather than the other items (prediction 5). However, even for that group, the difference was only statistically significant with a one-tailed test. A one-tailed test was appropriate because the prediction tested was unidirectional. These effects were not replicated in all the groups. More research is needed to identify the key variables involved in the design of the clips and whether, where appropriate, in practical design situations the effects can be amplified to achieve design objectives more dependably.

4.6. Hypothetical duplication and enhancing effects of adding a visual component

The findings from the present study suggest an interesting hypothesis concerning the ways in which multimedia designers in practice may use the visual component. In particular, Table 2 shows that group 2 did better than guessing when answering questions about clips that had only the visual component (no audio). Since the questions were based on the information in the audio, they could only have achieved this if the visual component contained information that effectively duplicated some of the information in the audio. We could therefore hypothesise that multimedia designers sometimes use the visual component to duplicate information in the audio, and that this can have positive value in terms of aiding memory for the information. We could refer to this as duplication of information.

Adding a visual component also seems sometimes to enhance the effectiveness of the audio. We could refer to this as an enhancing effect. Evidence for this comes from Table 1, which shows that group 1 scored better in the memory test when answering questions about clips that contained both the visual and the audio components than when answering questions about audio-only clips. Since the information needed to answer the questions was all in the audio component, this effect must have been due to the visual component in some way adding to the effectiveness of the audio in imprinting the information on the subjects' memory. Whether this enhancing effect comes from a simple duplication of the information or from something else cannot be determined from the present data. For example, one might speculate that designers sometimes use the visual component to complement the information in the audio, rather than simply duplicating it, and that the complementary information somehow helps subjects to remember what is in the audio. More research is needed to test these and other possibilities.

The effect of the visual component on remembering was not replicated in all groups. The reason for this is not clear from the present data. More research is needed. A first step should be to see if the effect can be replicated or whether it was a statistical anomaly.

4.7. Possible bias from the participants

Standard texts on psychology (e.g. Coolican 1998) typically identify a number of ways in which participants can bias the results of experiments, including a tendency they may sometimes show to deliberately or inadvertently adjust their behaviour to support what they believe to be the hypothesis being tested. In the present experiment, for example, it is possible that some or all of the subjects might have rated the items with a visual component as relatively more interesting because that is what they thought the experimenters expected. It is possible that the other ratings and estimates provided by the subjects were affected similarly. It is also possible that the subjects could have artificially boosted the benefit of a visual component in their results on the objective tests of memory, most easily by depressing their scores for the items that did not have a visual component.

If any such bias from the participants did amplify the benefits of adding a visual component, then that strengthens the conclusion that those benefits were not especially dramatic in the present experiment. Whatever effects there were, were relatively modest and not always easy to demonstrate.

4.8. Implications for designers and their clients

The present results support the hypothesis that adding a visual component to what would otherwise be an audio-only clip can add value, particularly in terms of how much information the clip is considered to contain, how interesting it is and, possibly, how confident users might feel in answering questions about it.

However, the results also suggest that, in practical situations, developers cannot take it for granted that there will always be value in other respects. For example, Table 1 shows that group 3 did less well on the memory test and reported that their attention wandered more when the visual component was added to the audio, although neither difference was statistically significant.

This suggests that developers and their clients need to give serious consideration to what exactly they are trying to achieve and whether the significant sums typically involved in producing audio-video clips could perhaps be better used on other aspects of the overall multimedia production. A corollary of this is that there may be some benefit in training multimedia developers in when and how to add vision to audio most effectively – an aspect of the 'media literacy' to which Nagy *et al*. (1999) refer. One might speculate that this might be particularly helpful where the developers have a stronger background in traditional software engineering than in the 'media'. In this connection, as Bobrowicz and Christie (2003) suggest, there may be a particular need for designers to have a better understanding of the role of multimedia montage and collage.

The findings from this study also underline the potential value of testing multimedia productions on prospective users. The alternative of relying exclusively on assumptions about how the productions will be received, based on the experience and good judgement of the development teams involved, may often not be good enough. It may be superficially attractive in minimising costs in the short term, but it is likely to be sub-optimal if it means design goals are not adequately achieved. Development teams will often be in a better position to draw on relevant past experience and to exercise good judgement when they have access to inputs from prospective users. This is a principle that is well recognised in the field of usability engineering (Dix *et al*. 2004) and lends itself naturally to production of multimedia applications (England and Finney 2002). It is also consistent with a movement towards greater user-orientation in the communication of information, even technical information (cf. Zachry *et al*. 2001). Such testing means involving prospective users at all stages, from concept to final product. Iterative prototyping and evaluation would form an important part of that wherever possible. The present findings underscore the importance of the evaluations using appropriate methods. For example, the present results suggest that prospective users may not always be very good at assessing how much they remember from clips. It follows that if a design objective is to facilitate learning of information presented, then objective tests of that should be included in the evaluations, rather than relying on users' estimates.

5. Conclusions

The present study supports the hypothesis that adding a visual component to what would otherwise be an audio-only clip can add value. It seems capable of increasing how much information the clip is considered to contain and how interesting it is considered to be. It may also help users to feel more confident in answering questions about it, although more research is needed to assess the replicability of that finding.

The study also illustrates that some effects cannot be taken for granted. Several predicted effects were observed for some groups but not for others. This suggests that developers and their clients need to consider carefully what they are aiming for and not take it for granted that they can achieve some effects easily.

The findings are also consistent with the view that involving prospective users in the development process at all stages might often be helpful in checking developers' assumptions and providing them with useful feedback on the extent to which they are achieving design objectives, but appropriate evaluation methods must be used.

References

Bobrowicz, A. and Christie, B., 2003. Usability and emergent properties of multimedia creations. Paper presented at the *Europrix Scholars Conference 2003, 'Information Society and Digital Media for All'*, 13–14 November 2003. Tampere, Finland.

Brinkman, W.P., Buil, V.P., Cullen, R., Gobits, R., and Van Nes, F.L., 2001. Design and evaluation of online multimedia maintenance manuals. *Behaviour & Information Technology*, 20 (1), 47–52.

Bruce, V., 1996. The role of the face in communication: implications for videophone design. *Interacting with Computers*, 8 (2), 166–176.

Bull, M., Christie, B., and Collyer, J., 1995. Multimedia in business. *Electrotechnology*, October/November, 18–20.

Carney, R.N. and Levin, J.R., 2002. Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14 (1), 5–26.

Coolican, H., 1998. Research methods. *In*: M. Eysenck, ed. *Psychology: An Integrated Approach*. Harlow, UK, Pearson Education Limited, 668–729. *Educational Psychology Review*, 14 (1), 5–26.

Dix, A., Finlay, J., Abowd, G.D., and Beale, R., 2004. *Human-Computer Interaction*, third edition. Harlow, UK, Pearson Education Limited.

England, E. and Finney, A., 2002. *Managing Multimedia: Project Management for Web and Convergent Media, Book 1 – People and Processes*, third edition. Harlow, UK, Pearson Education Limited.

Eysenck, M., 1998. Memory. *In*: M. Eysenck, ed. *Psychology: An Integrated Approach*. Harlow, UK, Pearson Education Limited, 167–204.

Faraday, P. and Sutcliffe, A., 1997. Designing effective multimedia presentations. *CHI 97*, 22–27 March 1997, 272–278.

Furnham, A., De-Siena, S., and Gunter, B., 2002. Children's and adults' recall of news stories in both print and audio-visual presentation modalities. *Applied Cognitive Psychology*, 16 (2), 191–210.

Hashmi, M.A. and Guvenli, T., 2001. Multimedia content on the web: problems and prospects. *Managerial Finance*, 27 (7), 34–41.

Matarazzo, G. and Sellen, A., 2000. The value of video in work at a distance: addition or distraction? *Behaviour & Information Technology*, 19 (5), 339–348.

Mayer, R.E., 2001. *Multimedia Learning*. Cambridge, Cambridge University Press.

Meisel, S., 1998. Videotypes: considerations for effective use of video in teaching and training. *Journal of Management Development*, 17 (4), 251–258.

Nagy, A., Collyer, J., Christie, B., and Southworth, K., 1999. Usability: a multimedia concept. Paper presented at the *Europrix 99 Conference, 'Understanding the European Content Industries – Towards a European Multimedia Economy'*, 19–20 November 1999. Tampere, Finland.

Panteli, N. and Dawson, P., 2001. Video conferencing meetings: changing patterns of business communication. *New Technology, Work and Employment*, 16 (2), 88–99.

Pardun, C.J. and Lamb, L., 1999. Corporate web sites in traditional print advertisements. *Internet Research: Electronic Networking Applications and Policy*, 9 (2), 93–99.

Records, H.A. and Olinsky, A., 1998. Dynamic financial imaging: using multimedia to measure the health of your business. *Industrial Management & Data Systems*, 98 (1), 17–22.

Schnotz, W., 2002. Commentary: Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, 14 (1), 101–120.

Shen, Q., Li, H., and Deng, Z., 2001. Intelligent multimedia for expertise-intensive training in construction. *Logistics Information Management*, 14 (5/6), 386–392.

Sweller, J., 1999. *Instructional Design in Technical Areas*. Camberwell, Australia, ACER Press.

Sweller, J. and Chandler, P., 1994. Why some material is difficult to learn. *Cognition and Instruction*, 12, 185–233.

Tiernan, S.L. and Grudin, J., 2000. *Fostering Engagement in Asynchronous Learning through Collaborative Multimedia Annotation*. Technical report MSR-TR-2000–91. Redmond, Washington, Microsoft Research, Microsoft Corporation.

Zachry, M., Faber, B.D., Cook, K.C., and Clark, D., 2001. The changing face of technical communication: New directions for the field in a new millennium. *In*: *Proceedings of SIGDOC '01*, 21–24 October 2001. Santa Fe, New Mexico, USA.