Pedagogic design guidelines for multimedia materials: a mismatch between intuitive practitioners and experimental researchers

Educational Media Production Training

Jack Koumi [jack.koumi@btinternet.com]

Abstract

This paper argues that pedagogic efficacy for multimedia packages cannot be achieved by experimental or by summative research in the absence of a comprehensive *pedagogical screenwriting* framework. Following a summary of relevant literature, such a framework is offered, consisting of micro-level design and development guidelines. These guidelines concentrate on achieving pedagogic synergy between audio commentary and visual elements. The framework is grounded in the author's experience of producing multimedia packages at the UK Open University.

Introduction

This paper offers micro-level design guidelines for pedagogic harmony between sound and images in multimedia packages. These guidelines are compared with design recommendations in the research literature. It is argued that such recommendations have minimal value for practitioners because they address the macro-level. To be of practical use, the research needs to derive from micro-level design principles that are tacitly employed by practitioners.

Van Merriënboer (2001) notes that little is known about the optimal combination of audio or speech, screen texts, and illustrations in pictures or video.

In fact, some substantial papers do exist, written by educational technologists such as Laurillard and Taylor at the UK Open University. These address several detailed design techniques, which appear below, but mainly they discuss over-arching, macro-level questions, such as how learners might cope without a fixed linear narrative (Laurillard, 1998; Laurillard et al, 2000) and an *analytical framework for describing multimedia learning systems* (Taylor, Sumner and Law, 1997).

However, for the practitioner who is trying to design a pedagogically effective package, the literature is of little help. There appears to be no published *comprehensive* framework of micro-level design principles for optimal integration of visuals and audio. This is despite the many investigations into the use of audio commentary in multimedia presentations. Some of these investigations are summarised below, exemplifying the mixed results in the comparison of screen text with audio commentary.

Following this summary, the major part of this paper presents a framework of design guidelines for multimedia packages. These guidelines are in the form of practicable, micro-level pedagogic design principles, such as

- there are occasions when the audio commentary should come first, in order to prepare the viewer for the pictures, such as, *In the next animation, concentrate on the arms of the spinning skater <ANIMATION STARTS WITH SKATER'S ARMS HELD WIDE, THEN PULLED IN>*

The framework has been compiled from the practices of designers of multimedia packages at the UK Open University. It incorporates an abundance of *practitioners' knowledge* regarding pedagogic design of audio commentary and graphic build-up. The width and depth of the framework offers a substantial basis for future investigations – a set of design guidelines that can generate fruitful hypotheses.

The literature relating visuals and audio commentary

Tabbers, Martens and Van Merriënboer (2001) report several recent studies by Moreno, Mayer, and others, in which multimedia presentations consisted of pictorial information and explanatory text. Many of these studies demonstrated the superiority of audio text (spoken commentary) over visual, on-screen text. In various experiments learners in the audio condition spent less time in subsequent problem solving, attained higher test scores and reported less mental effort. The investigators attributed these results to the *modality effect.* This presupposes *dual coding*, whereby auditory and visual inputs can be processed simultaneously in *working memory*, thereby leaving extra capacity for the learning process.

In their own study, Tabbers et al (ibid) presented diagrams plus audio commentary to one group, but to a second group they replaced the audio commentary with identical visual text, on screen for the same duration. They found that the audio group achieved higher learning scores. However, when two other groups spent as much time as they liked on the same materials, the superiority of the audio condition disappeared. The authors conclude that the purported modality effect of earlier studies might be accounted for in terms of lack of time rather than lack of memory resources. (Mind you, the students in the visual text condition had to spend longer on task to achieve their comparable scores, so the audio condition could still claim superior efficiency)

Others have found that addition of audio need not be beneficial to learning. Beccue, Vila and Whitley (2001) added an audio component to an existing multimedia package. The audio was a conversational version of a printed lab manual that college students could read in advance. The improvement in learning scores was not statistically significant. Many students suggested that the audio imposed a slower pace than they were used to. The authors theorized that the pace set by the audio might be helpful for slow learners and detrimental to fast learners

Kalyuga (2000) observed a similar effect, finding that *novices* performed better with a diagram plus audio than with a diagram-only format. However, the reverse was found for *experienced learners*.

In another experiment, Kalyuga (ibid) found that audio commentary did indeed result in better learning, but only when the identical visual text was absent. Specifically, a diagram was explained in three different ways: visual text, audio text, visual text presented simultaneously with identical audio text. The visual-plus-audio group achieved much lower scores than the audio-only.

Kalyuga's interpretation of this result was that working memory was overloaded by the necessity to relate corresponding elements of visual and auditory content, thus interfering with learning. He concluded that the elimination of a redundant visual source of information was beneficial.

However, this interpretation should predict that elimination of a redundant *audio* source would also be beneficial, i.e. that the visual group would learn better than the visual plus audio group. In fact, the result was slightly in the opposite direction, which also meant that the audio only group learned much better than the visual-only group. Hence a more convincing explanation is a *split attention* effect. In the visual-only condition, students had to split visual attention between the diagram and the visual text. This imposes a greater cognitive load than the audio-only condition, in which students had only one thing to look at (the diagram) while listening simultaneously to a spoken description.

Moreno and Mayer (2000) presented an animation accompanied by either audio text or visual text also found a strong split-attention effect, which they express as a Split-Attention Principle:

Students learn better when the instructional material does not require them to split their attention between multiple sources of mutually referring information (in their experiment, the information in visual text referred to the information in the animated diagrams and vice-versa)

In a refinement of these experiments, Tabbers, Martens and Van Merriënboer (2000) compared two strategies for decreasing cognitive load of multimedia instructions: *preventing split-attention* (preventing visual search by adding visual cues) or *presenting text as audio* (replacing screen text with audio commentary). They found that students who received visual cues scored higher on reproduction tests. However, the modality effect was opposite to that expected, in that visual text resulted in higher scores than audio commentary.

The authors advanced some speculative reasons for this reversal of previous findings:

- students reported expending significantly greater mental effort in the visual text condition. Whatever the reason for the greater effort (possibly that reading visual text is more learner-active than listening to audio text), it could have resulted in deeper processing and hence better learning.
- students could choose to replay the audio text segments (and to re-read the visual text). However, in both conditions, it is likely that students had partly understood the texts on first listening (or first reading). Hence, in the visual text condition, students who re-read the text could skip any visual text that they had already understood, whereas in the audio condition, students who re-listened would be forced to process some redundant auditory information.

These are reasonable conjectures for superior learning in the visual text condition. A third likely reason (which does not conflict with the two conjectures) was the complexity of the task. Students studied how to design a blueprint for training in complex skills, based on Van Merriënboer's *Four Component Instructional Design* model. The task is certainly complex. It necessitates self-paced, head-down, concentrated study of complicated diagrams and relationships (students were allowed an hour to work through the multimedia learning task). As argued by Koumi (1994), such tasks cannot easily be supported by audio commentary, because this is a time-based (transient) medium. Instead, what's needed is a static (printed) set of guidelines that students can revisit repeatedly while they carry out intensive, self-paced study of the diagrams.

The above arguments may throw some light on the various conflicting results. However, there may be more fundamental reasons for the inconsistencies, as follows.

Hede (2002) notes that the conflicting results are not surprising, considering the myriad of contingent factors that have been shown to moderate multimedia effects, including:

- nature of visual and audio input
- positive and negative modality effects (dual coding vs. interference)
- interactivity and cognitive engagement
- cognitive overload

A related factor is the following. The cited experimental studies, in manipulating the format of a multimedia package, may have introduced debilitating distortions into a previously harmonious pedagogical design. If so, the inconsistent results might be artefacts of design distortions. Moreover, the experimenters cannot easily control for these distortions, because to date, there are no published micro-level design guidelines that focus on harmony/synergy. This paper aims to provide such a framework.

The provenance of the Design Framework

UK Open University multimedia packages are typically produced over several script conferences by a team of experienced teachers who know their target audience well. For such a team, permitted sufficient thinking-time, the ensuing learning material is based upon several lifetimes of teaching experience.

Successive script conferences build creative momentum in which the critical analysis becomes progressively deeper. Effectively, the team is carrying out a whole *series* of developmental re-evaluations, as *thought experiments*, each member of the team repeatedly taking on the role of a hypothetical student. In addition, many of these design teams include an educational technologist, who contributes research experience and knowledge of current learning theories. Over time, the team will have developed a tacit, intuitive design model.

This paper seeks to pull together these tacit design models and make them explicit, in the form of the framework below. No doubt the framework owes a debt to learning theories that have permeated the collective psyche of audiovisual practitioners. However, such derivation has not been a conscious pursuit. The framework has been successively refined through the author's appraisal and co-authorship of UK Open University multimedia packages and those of other institutions. Critical comments regarding this first published attempt will be welcomed.

The Design Framework

A multimedia package might include video clips containing their own commentary. The screenwriting principles for designing video commentary are beyond the scope of this paper. Chapters 5 and 6 of a forthcoming book by Koumi (2006) provide a framework of such principles. A pre-cursor to these chapters is the paper by Koumi (1991).

However, when the rest of the multimedia package also contains an audio commentary, there are further screenwriting principles to consider. These principles/guidelines are summarised in section 4 below. This section is preceded by some practical points, regarding graphics in section 1 and regarding the production/development process in sections 2 and 3. The full framework is longer, including examples and subsidiary principles. It forms Chapter 8 in Koumi (2006).

1. The visuals

The visuals can be equations, printed text, (both often built up line by line from top to bottom of the screen), diagrams, animations, video. Usually, the screen would be divided into sections, e.g. video on the left, text on the right.

In all cases, the visuals could be *pure source material*, which the audio teaches about, or can *incorporate their own teaching*, in the form of visual text, which the audio elaborates on. Or, there could be a mixture; that is, there would be some screen text accompanying the visuals, giving an outline explanation, and the audio would elaborate on the outline.

Some of the text would be in the form of interactive dialogue boxes whereby students carry out activities, inputting their own text in response to questions.

2. How to prepare for the production

2.1. Consider / Specify:

- the learning tasks in this week's syllabus that would benefit from a multimedia package
- the intended learning outcomes
- learning context
- target audience characteristics

2.2. Decide on software and delivery platform

Which programming environment should be used? Should the package be delivered on a CD-ROM, via the Web, or both? Should there be links to commercial software? For an informative discussion, see Taylor et al (ibid).

2.3. Decide on type of visual materials

For example, is video needed? Should the graphics be in the form of 2D diagrams that are built up in stages or should the diagrams be 3D, with full single-frame animation?

2.4. Compose an Outline, on paper, of the multimedia "screens"

A **screen** consists of a sequence of visuals/graphics that develops autonomously over time. For example, a title might appear at the top, followed one second later by an equation on the left. This might be followed, after a sentence of narration (audio commentary), by a second equation, followed two seconds later by a phrase of printed text, culminating with another sentence of narration. The next **screen** of graphics starts when the student elects to move on, e.g. by pressing *NEXT*.

2.5. Design screens to constitute a full outline of the topic

It is a good idea for the screens alone to constitute a full outline of the content, sufficient to enable busy colleagues to evaluate your design without listening to the commentary.

3. The Production

3.1. Record a first draft audio guide-track and program the visuals to appear at particular words.

3.2. Finalise the *screens* and get a print-out. This finalisation will require several iterations (and the programming modifications that are required to implement each draft will rarely produce precisely what you envisaged).

3.3. Rehearse then record the final sound track. On both occasions, someone should take on the role of the student, looking at a printout of the screens while hearing the speaker. There will be one or more occasions when the match between the commentary and the graphics can be improved by changing one or the other.

3.4. Digitise into individual files, one for each screen. Lay onto the multimedia package audio-line

3.5. Adjust the picture build-up so that it is geared to the final sound-track (until this stage, the pictures were geared to the guide-track).

4. Pedagogic guidelines for screen/audio design for multimedia

The guidelines are divided into several categories:

- Navigational guidance and student control
- Use of language
- Layout of the screen
- Relationship of screen text to audio commentary
- The visuals and the commentary should reinforce each other
- Interactive elements
- Educational narrative – judiciously balance *effective exposition* by the teacher, against *independent exploration* by the student

Navigational guidance and student control

4.1. Start with a *Contents page* from which learners can access the different sections (normally in whatever

order they wish). The *Contents* page should record where students have been (e.g. the title of a section should get a marker once a learner has accessed it).

4.2. In each screen, an *audio-bar* should move to indicate how far the audio file for that screen has progressed.

4.3. Each audio file corresponds to a screen. That is, the graphics build-up finishes at or before the end of the audio file. Hence students have a visual indication of the progress of the screen information.

4.4. As noted by Taylor et al (1997), when students are revisiting a *screen*, they do not always want to listen to the audio track. This can also be true of experienced students visiting for the first time, as discussed by Beccue et al (2001) and Kalyuga (2000). It could also be true of busy colleagues who are formatively evaluating your design, as discussed in item 2.4.

User choice of whether to hear the commentary can be achieved by including a *skip* button (next to the *audio-bar*), with which learners can jump to the end of the current audio file. This would also skip past the graphics build-up, jumping straight to the full-screen graphics.

4.5. More generally, students should be free to skip to any section/chapter of the package, in any order they wish. They might thereby lose the narrative. However, the contents page (or map), tells them the teacher's intended structure.

Use of language

4.6. Long sentences should be avoided. They might exceed the listener's memory span and they often contain conditional clauses, which are difficult to bear in mind.

4.7. The narration is audio not print, so write *conversational speech*, to be spoken and listened to, *not* to be read. Here is one way to achieve this:

- draft out the *screens* before writing a commentary script
- *speak to* this draft, in a conversational style, straight into a tape recorder
- listen to the tape and transcribe it, then work on the transcript to improve it (but keep it conversational)

Layout of the screen

4.8. Students cannot easily process dense visual layout while listening to commentary. In particular, concerning visual text, a rule of thumb is to use only 25% of normal print density.

4.9. Even when the text is sparse, the standard technique is to *develop* a screen of graphics line by line. However, such piecemeal development of graphics may leave learners feeling *blinkered*. So, occasionally, you may feel it is appropriate to reveal two or three lines of text, allowing learners to choose whether or not to read ahead.

Relationship of screen text to audio commentary

4.10. The first question that arises – why include explanatory screen text at all? Could the audio commentary not suffice? One reason for *succinct* items of screen text, is that these can serve as visual reference points, which anchor attention. They can also prevent overloading auditory memory.

4.11. Many multimedia packages with audio commentary present *identical* visual text simultaneously – a practise that is emulated by most of the comparison studies discussed earlier. However, literate students can read faster than you can speak. Consequently, the asynchronous semantic processing of the two sources causes mutual interference. This can be avoided if the screen text is a judicious précis of the audio commentary, not a duplicate. (Sufficient text to enable subject matter experts to understand the content without listening to the narration – see 2.4 and 4.4)

4.12. When the audio narration is précised by on-screen text, students will search the text in an attempt to track what they are hearing. If there is more than one line of text (as recommended in 4.9) this tracking could be difficult, which would disrupt students' understanding. Hence the on-screen text should *reproduce key words* of the narration, and the narrator should speak these key words verbatim rather than paraphrasing them.

4.13. It is useful to include a transcript of the audio commentary that students can access for each screen by clicking on a *SCRIPT* button. Then, for example, if students are revising by skimming/browsing through the screens, they can have immediate access to the whole commentary relating to each complete screen. In any case the transcript is essential for deaf students.

4.14. However, as recommended in 2.5, you should design screens to be a full outline of the topic, sufficient for your busy colleagues to grasp without having to listen through the audio commentary. With such a design, few students would need to access the transcript during their revision, because if the screens are sufficient for your expert colleagues, they should also be sufficient for most students when they revise – because they have previously heard the explanatory narration.

The visuals and the commentary should reinforce each other

4.15. In many situations, the words should synchronise with the corresponding visuals, for example when the narration refers to parts of a diagram that are being successively highlighted. However, there are many situations in which the words should precede or follow the corresponding visuals, as follows.

4.16. Make teaching points about a visual when students are looking at it, and not in a wordy introduction while they are looking at the previous visual. This applies particularly in mathematics - the pictures need to precede the words when the pictures are mathematical expressions that are difficult to listen to unless they can be seen.

4.17. However, there are occasions when the words should come first, in order to prepare the viewer for the pictures, such as, *In the next animation, concentrate on the arms of the spinning skater. <ANIMATION STARTS WITH SKATER'S ARMS HELD WIDE, THEN PULLED IN>*

4.18. Give students enough time to digest the visuals. A notorious error is to position the words at the beginning of the audio file rather than preceding the words with a suitable pause, say 2 seconds.

4.19. Indicate clearly where to look on the screen, either by a visual cue such as highlighting an item when it is mentioned or by a verbal cue such as, *notice the top of the diagram*, or both.

Interactive elements

4.20. Whenever students carry out an activity, they should be able to keep a record, inside the package, e.g. typing into a *notepad*, rather than on a scrap of paper, as exemplified in the Homer package (see Laurillard, 1998). This serves to preserve the narrative that has been *co-authored* by the package and the student.

4.21. The package should provide appropriate *scaffolding* for student activities. For example, if a student types an incorrect answer into a dialogue box, there should be an option to get one or more hints then try again, and eventually to be told the correct answer. If there is no "correct" answer (as in open questions), the package should still afford feedback, in the form of a model answer. Laurillard (1998) recommends withholding the model answer until the student has made a sizeable attempt.

Educational narrative: judiciously balance structured exposition by the teacher, against independent exploration by the student

The efficacy of narrative structure has been proposed by many writers, such as Gudmundsdottir (1995), Gibson (1996), Laurillard (1998) and Laurillard et al (2000). The guidelines below are intended to structure each *chapter* of the narrative. They are treated only in outline, adapted from the principles for pedagogic video design, in Koumi (1991 and 2006).

4.22. **Signpost**: Clearly indicate where the chapter is going, what is happening next, why it is happening, what to look out for.

4.23. **Facilitate concentration**: e.g. short pauses for contemplation, encourage prediction.

4.24. **Encourage / enable constructive learning** e.g.

- concretise: that is, relate to (hence *activate*) students' previous knowledge
- support/scaffold the learner's construction of knowledge, e.g., concerning problem-solving, Merrill (2002) and Van Merriënboer (2001) recommend several techniques, such as
- show a worked example of a similar problem before asking students to solve a problem
- begin with a simple *half-way-there* problem
- divide the problem into sub-problems
- pose a sequence of successively more difficult problems
- the type and extent of scaffolding depends on the ability of the learner in relation to the material – and the scaffolding should be withdrawn little by little as learners progress.

4.25. **Elucidate**: moderate the load, pace and depth, maximise clarity

4.26. **Reinforce**: e.g. give more than one example of a concept, use *comparison* and *contrast*, ensure *synergy* between commentary and images (e.g. as in 4.15 to 4.19).

4.27. **Consolidation** of learning could be achieved through students solving end-of-chapter problems and referring to model answers.

How can narrative coherence be salvaged?

Gibson (1996) predicted correctly that multimedia authors would continue to incorporate many different layers in a single work and to enable *a wide variety of paths through (the material)*, so that learners can piece together their own, idiosyncratic stories. The design considerations for such multi-layered works are rather complex, in that all of the diverse possible narrative structures of the viewers need to be borne in mind. One way to restrict this indeterminacy would be to discourage students from progressing from chapter to chapter until they make a sizeable attempt at the end-of-chapter consolidation activity. Also, the contents page (or map), tells student's the teacher's intended structure

Other strategies for preserving the narrative are discussed by Laurillard et al (2000), including:

- a clear statement and repeated reminders of an overall goal – to support a task-related plan
- a model answer – to motivate reflection on their conceptions

Comment. Concerning the statement of the goal and reminders of the goal, these can be progressively more specific as learners progress through the package and become more able to understand the language and what is expected of them.

The mismatch between intuitive practitioners and experimental researchers

Two types of research papers were outlined at the beginning of this paper. One type reported on experimental studies of individual audio and visual variables. The other type, by UK Open University writers, were summative studies that dealt primarily with macro-level design issues. How do these two sources relate to the above framework of micro-level design guidelines?

The micro-level guidelines in relation to summative, macro-level studies

The guidelines might serve to add flesh to the further development of the over-arching issues espoused in the macro-level papers by Laurillard, Taylor and others. In return, such issues need to be borne in mind for future development of design guidelines.

The micro-level guidelines in relation to experimental studies

The guidelines derive from practitioners. They are more detailed than the levels of investigation carried out in the experimental studies described earlier. This discordance is natural. The variables that can be investigated using a scientifically acceptable experimental study are simpler than the complex integration

of design principles that must be used by practitioners.

On the other hand, these design principles are intuitive and have not been studied scientifically. The framework of design guidelines is offered as a fledgling design theory for researchers to investigate the practitioners' intuitions. It would be heartening if this paper could start an iterative process whereby researchers and practitioners collaborate to improve the design of multimedia packages.

Currently, it appears that there is no widespread collaboration between practitioners and researchers. Instead, the aforementioned experimental studies build on theoretical interpretations of previous experiments, such as those compiled by Moreno and Mayer (2000). Based on these results and on various learning theories, the authors propose a cognitive theory of multimedia learning that comprises six principles of instructional design:

- Split-Attention
- Modality
- Redundancy
- Spatial Contiguity
- Temporal Contiguity
- Coherence

These six principles exemplify the mismatch between the research literature and the concerns of practitioners.

The Split-Attention principle was discussed earlier (visual attention detrimentally split between screen text and corresponding diagrams). The principle is intuitively reasonable. Note however that it leads to the either/or recommendation that audio text is always superior to screen text. So there is no conception of a judicious combination of the two, as recommended in items 4.15 to 4.19 – namely that screen text should be a judicious key-word précis of the narration, serving as visual mnemonic and an anchor for the narration. This does leave open the danger of splitting visual attention, but with sparse screen text, the effect should not be too detrimental.

The Modality principle asserts that

Students learn better when the verbal information is presented auditorily as speech than visually as on-screen text both for concurrent and sequential presentations

It was noted earlier that Tabbers et al (2000) reported a counterexample in the case of a complex task that required self-paced reflection of the on-screen text.

Furthermore, note the surprising *sequential* condition, in addition to the usual *concurrent* presentation of explanatory text. Moreno and Mayer actually tested the effect of presenting the whole text *before* the whole animation and also *after* the whole animation (and found that audio text was superior in both conditions, as well as when presented concurrently).

The purpose was to determine whether the superiority of audio text was a memory capacity effect (screen text and diagrams overloading visual working memory) rather than a split-attention effect (insufficient attention paid, in the concurrent presentation, to either or both visual components – screen-text and diagrams).

Such extreme manipulation of the variables might possibly help to build a learning theory, but is of little use to the practitioner. Of what use is it to know *for sequential presentation* that the auditory condition is superior to the visual? If an animation needs to be complemented by audio commentary (true in most cases), there is no point in delaying the commentary rather than synchronising it. No multimedia designer would contemplate such a design because it would severely limit any *synergy* between commentary and diagrams.

This leads to a serious point. Creating synergy between diagrams and synchronised audio commentary is not a trivial endeavour. A bad designer could accidentally fabricate a package in which the composition and pacing of the audio commentary actually clashed with the concurrent diagrams, hence *interfering* with learning. With such a disharmonious design, the dissonance might be reduced by separating the diagrams from the commentary – presenting diagrams and commentary sequentially rather than concurrently. The sequential package might still fail, but not quite as badly.

Much better to accept the challenge of creating synergy in a concurrent presentation. Techniques for achieving such synergy are described in items 4.15 to 4.19.

The Spatial Contiguity principle asserts that

Students learn better when on-screen text and visual materials are physically integrated rather than separated

This is not surprising, but the mistake of spatially separating text from diagrams is quite common, so the principle is worth stating.

In another sense, the principle is rendered redundant by another finding from the experiment that supported it. It was also found that replacing the on-screen text by audio narration produced even better learning (hence re-confirming the modality principle). This would render the Spatial Contiguity principle useless, since the recommendation would have to be that on-screen text should be deleted altogether.

However, the authors used text that duplicated the audio rather than being a judicious précis. It was argued above that key words of the audio narration should be presented as on-screen text, thereby *reinforcing* and *anchoring* the narration. In such a design, the contiguity principle still has currency – that there should be facilitative placing of text and diagrams.

The Temporal Contiguity principle asserts that

Students learn better when verbal and visual materials are temporally synchronized rather than separated in time.

However, this principle is a macro-level guideline that cannot help the practicing multimedia designer. In fact, the authors and others have demonstrated that the principle does not hold unless the temporal separation is considerable (e.g. when a large chunk of animation is preceded by the whole narration). Compare this principle with the micro-level guidelines 4.15 to 4.17 above, which exemplify the fine judgments of pacing and sequence of the intuitive designer who *gets inside the head of the learner*.

Again, consider the Redundancy Principle, which asserts that

Students learn better from animation and narration than from animation, narration, and text if the visual information is presented simultaneously to the verbal information

This principle assumes that the text is identical, word for word, with the narration. This approach was rejected a priori by UK OU designers, who surmised that simultaneous reading and listening would be uncomfortable (item 4.11 elucidates this rationale). Instead, the concern of OU designers is the subtle issue of how succinct should be the text so that it anchors the narration (see 4.9, 4.10 and 4.12) but does not interfere with it.

The Coherence principle asserts that

Students learn better when extraneous material is excluded rather than included in multimedia explanations

This principle was supported in an experiment by Moreno and Mayer (ibid) in which learning was significantly worse when music was added to an animation with narration. This effect relates to one with a more evocative name, the *seductive-augmentation effect* – defined by Thalheimer (2004) as a negative effect on *learning base material* when the presentation is augmented by interesting but inessential text, sounds or visuals (the augmentation *seduces* the learner's attention/processing away from the essential items). Thalheimer reviewed 24 research comparisons, 16 of which showed that adding interesting items hurt learning, 7 showed no difference and one showed a learning increment.

Intuitively we feel that making the presentation interesting is a good idea, because it engages the learner. But how does this jibe with the above results? Here are two possibilities:

In the UK Open University, producers would typically spend hours choosing music that was appropriate to the mood of the story. Typically, even after sifting through printed descriptions of music tracks and discarding many choices, at least 80% of the music tracks chosen for consideration were found to *jar* with the storyline. In any case, music was normally only played when there was a deliberate pause in the commentary, designed to allow viewers to reflect on the pictures. It is clear that the above experimenters did not follow these provisos, so it is not surprising that their music interfered with learning.

A second interpretation of the results is speculative but intriguing. All 16 experiments showing a significant negative effect involved very short learning tasks, average 4 minutes. An interesting conjecture by Thalheimer is that for longer tasks, in which attention might flag, adding interesting elements to sustain attention might have a positive effect *on balance*. That is, if the seductive augmentations do indeed cause a learning decrement (say 20%), this may be the price we have to pay to keep learners attentive for longer periods. That is, if seductive augmentations really do distract, the negative effect may be more than compensated by their *sustained-attention* effect.

Recommendations for future research and design development

To conclude, all six principles recommended by Moreno and Mayer (2000) are pitched at a macro level that may be suitable for theory building but that only skims the surface of the detailed design concerns of the practitioner.

A diligent search of the literature has failed to uncover any more practicable design principles. Admittedly, the literature could serve as a useful backdrop for the practitioner. However, value is more likely to be obtained in the reverse direction. Namely, before we try to derive design principles based on macro-level learning theories, and refine them through experimental studies, we would be better advised to start from experienced teachers' intuitive, micro-level design guidelines, and progress in the opposite direction, namely

micro-level design principles, as espoused above (a fledgling design theory) → *experimental studies* → *refined theory/design principles* → *...*

Each of the proposed design principles could generate questions to be investigated. For example

- Regarding guideline 4.10, how precisely should the screen text be phrased, in relation to the audio commentary, so that it anchors the commentary rather than interfering with it?
- Regarding guideline 4.14, just *how* sparse should the screen text be? Should the screens by themselves (without the audio) be just barely comprehensible to a really top expert in the subject matter, or should they constitute a more substantial outline of the content?

Note that these questions concern the nature of the visual text rather than whether it is present or absent. This illustrates the philosophical conflict between the experimental studies and the design guidelines in this paper. The guidelines aim to integrate narration and visuals in order to achieve optimum synergy between these two constituents. In contrast, the aforementioned experimental studies manipulate the two constituents separately, thereby compromising their synergy. A good media designer would re-script the narration if denied harmonious visuals and would re-draught the visuals if denied harmonious narration. The above experimental studies could not countenance any such reconstruction of audiovisual synergy because this would have defeated their manipulation of the separate constituents!

Collaboration between researchers and practitioners would have a much increased chance of being productive if the investigations compared different ways of *integrating* narration and visuals – different composite designs aimed at optimum audiovisual synergy – rather than trying to unpick the two constituents and manipulate them separately.

References

[1] Beccue B, Vila J and Whitley L K (2001) *The effects of adding audio instructions to a multimedia computer based training environment*, Journal of Educational Multimedia and Hypermedia (2001) 10(1), 47-67 http://www.aace.org/dl/files/JEMH/jemh-10-01-47.pdf

[2] Gibson S (1996) *Is all coherence gone? The role of narrative in Web design,* Interpersonal Computing and Technology, 4, 2, pp. 7-26 http://www.helsinki.fi/science/optek/1996/n2/gibson.txt

[3] Gudmundsdottir S (1995) *The Narrative Nature of Pedagogical Content Knowledge*, in H. McEwan and K. Egan, *Narrative in teaching, learning and research.*New York: Teachers College, 24-38 http://www.sv.ntnu.no/ped/sigrun/publikasjoner/PCKNARR.html

[4] Hede A (2002) *An integrated model of multimedia effects on learning*, Journal of educational multimedia and hypermedia (2002) 11(2), 177-191 http://www.aace.org/dl/files/JEMH/JEMH112177.pdf

[5] Kalyuga S (2000) *When using sound with a text or picture is not beneficial for learning*, Australian Journal of Educational Technology 2000, 16(2), 161-172 http://www.ascilite.org.au/ajet/ajet16/kalyuga.html

[6] Koumi J (1991) *Narrative screenwriting for educational TV,* Journal of Educational Television, 17, 3, 1991

[7] Koumi, J (1994), *Media comparison and deployment: A practitioner's view*, British Journal of Educational Technology, vol. 25, no. 1, pp. 41-57

[8] Koumi (1995) *Building good quality in, rather than inspecting bad quality out,* Chapter 31in (Ed. Lockwood F (Ed) Open and Distance Learning today, London, Routledge

[9] Koumi J (2006), *Designing Video and Multimedia for Open and Flexible Learning,* Routledge Falmer

[10] Laurillard D M (1998), *Multimedia and the Learner's Experience of Narrative,* Computers and Education, 31 (2) 229-242

[11] Laurillard D M, Stratfold M, Luckin R, Plowman L and Taylor J (2000), *Affordances for learning in a non-linear narrative medium,* Journal of Interactive Media in Education, 2000 (2) http://www-jime.open.ac.uk/00/2/laurillard-00-2-paper.html

[12] Moreno R and Mayer R E (2000), *A Learner-Centered Approach to Multimedia Explanations: Deriving Instructional Design Principles from Cognitive Theory*, Interactive multimedia electronic journal of computer- enhanced learning 2000 Vol. 2 No. 2 http://imej.wfu.edu/articles/2000/2/05/index.asp

[13] Merrill, M D (2002), *First Principles of Instruction.* Paper in preparation, accessed from http://www.id2.usu.edu/Papers/5FirstPrinciples.PDF

[14] Tabbers H K, Martens, R L and Van Merriënboer, J J G (2000), *Multimedia instructions and cognitive Load Theory: split attention and modality effects*, paper presented at the AECT 2000 in Long Beach, California. http://www.ou.nl/Docs/Expertise/OTEC/Publicaties/huib%20tabbers/AECT2000%20Huib%20Tabbers.pdf

[15] Tabbers H K, Martens, R L and Van Merriënboer, J J G (2001), *The modality effect in multimedia instructions*, Annual conference of the Cognitive Science Society, 2001 http://www.hcrc.ed.ac.uk/cogsci2001/pdf-files/1024.pdf

[16] Taylor J, Sumner T and Law, A (1997), *Talking about multimedia: a layered design framework*, Journal of Educational Media, 23(2/3) 215-241

[17] Van Merriënboer Jeroen J G (2001) *Instructional Design for Competency-based Learning* Interactive Educational Multimedia, No. 3 (October 2001) pp. 12-26 http://www.ub.es/multimedia/iem/ then click *Past Numbers* then click *IEM, 2001, 3* then click title of paper