

Decisions on the combinations of different multimedia types

1. Theoretical background

Multiple and Single Channel Communication

Principles from the multiple and single channel communication theories are important and closely related to multimedia design because when exposed to information delivered through different multimedia types the user processes it through different perception channels. Knowledge on how this information is processed and the effect it has on the recipient is important because then the information can be presented in a way that will support and ultimately lead to an optimal...

Multiple channel communication is generally considered to involve a continuous simultaneous presentation of information over two or more channels. The basic idea behind this mode of presentation is that the additional information input through different sensory channels (i.e., sight, sound, touch, etc.) will provide additional stimuli reinforcement which in turn will assist in organising and structuring incoming information. (Dwyer, 1978).

On the other hand single channel communication theory is based on the limited amount of information a person can process. Thus the increased number of stimuli can impose severe limitations on the processing capabilities of the brain

Mayer (1999) proposed a theory of multimedia learning that incorporates three important themes from existing theories of human cognitive processing (Baddeley 1992; Chandler & Sweller 1991; Paivio 1986):

- Humans possess two information processing channels, one for visually-presented material and one for auditory-presented material
- The processing capacity (or working memory capacity) of each channel is limited
- Active processing involves selecting relevant visual and verbal information, organizing the material into coherent mental models, and integrating between visual and verbal representations as well as existing knowledge from long-term memory

As Mayer incorporates theories describing aspects of human cognitive processing that are characteristic not only for students but are valid for any user, his model of multimedia learning from animation and narration can be generalized for any situation in addition to the learning situation he is addressing.

Based on the above mentioned three themes Mayer constructs a dual channel model of multimedia learning from animation and narration. Figure 1 summarises Mayer's model.

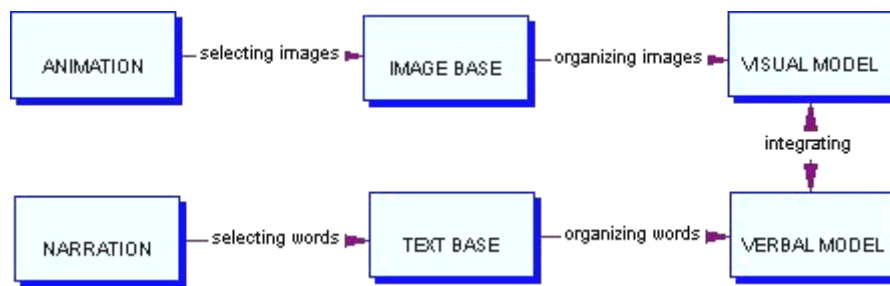


Figure 1. A dual channel model of multimedia learning

The boxes in the figure refer to mental representations and the arrows refer to cognitive processes. The model first presents the incoming pictorial information (such as an animation) and incoming verbal information (such as a narration), represented by boxes labelled animation and narration, respectively. By paying attention in the visual channel (represented by the arrow labelled selecting images), the learner is able to select some of the incoming animation for further analysis (represented as the box labelled image base). By paying attention in the auditory channel (represented by the arrow labelled selecting words), the learner is able to select some of the incoming narration for further analysis (represented as the box labelled text base). Next, in an attempt to build internal connections among the selected images (represented by the arrow labelled organizing images) the learner constructs a coherent representation of the images (represented by the arrow labelled visually based model). Similarly the learner constructs his/her verbally-based model. Finally, guided by relevant prior knowledge, the learner attempts to build external connections between the visually-based and the verbally-based models (represented by the arrow labelled integrating).

Based on his theory of multimedia learning Mayer defined the following principles for the design of multimedia instructions. Each principle is based on multiple empirical tests carried out at the Santa Barbara lab as well as a cognitive theory of multimedia design. The principles are:

- Multimedia principle it is better to present an explanation in words and pictures than solely in words
- Contiguity principle when giving an explanation, present corresponding words and pictures near rather than far from each other or simultaneously rather than successively
- Coherence principle when giving an explanation, exclude rather than include extraneous words, pictures, and background sounds
- Modality principle when giving an explanation, present words as narration rather than as on-screen text

Furthermore, in a series of experiments Mayer proved the relationship between the modality principle and split attention effect. When users are exposed to the combination of picture and written text, as both modalities are perceived through the visual channel, the effect of splitting attention occurs. This effect requires more mental effort and time to integrate the information coming from the two sources. On

the other hand such presentation allows easier and faster revisiting of those visual and verbal parts of the presentation which are related to each other

In contrast when picture and spoken text are combined they can be perceived simultaneously, no split-attention occurs and the information can be processed faster. But then later attempts to strengthen the integration between the already built visual and verbal model by revisiting corresponding parts between the picture and the text are more difficult as it is more time consuming to find a required segment of a spoken text.

2. Decisions for combining multimedia items

The following section attempts to identify whether the above mentioned principles are kept with the current multimedia presentation generation realized by Cuypers and how can Cuypers be improved in order to incorporate violated principles or those which hasn't been considered at all.

Multimedia principle

I assume that this rule is fulfilled as Cuypers basically retrieves all the media types available for a given query.

Contiguity principle

Mayer's contiguity principle sets spatial and temporal rules for combining textual and visual information. Spatially the two media types should be as close as possible, temporally they need to be simultaneously presented. The presentation generated by Cuypers is in alignment with the contiguity principle. Spatially the related media items are placed close to each other and temporally they are played simultaneously.

Coherence principle

I assume that Cuypers generates multimedia presentations with constituents relevant to the presentation topic hence the presentation will not include extraneous words and pictures.

Among the experiments that Mayer (1999) carried out to empirically support his coherence principle one in particular might raise interesting points to be considered in order to improve the generated presentations. The experiment shows outperformance of subjects exposed to concurrent narration and animation compared to ones who received the same narration and animation but with additional music and environmental sounds. Thus the following presentation rule can be formulated: If auditory and pictorial information are combined where the auditory information is narration explaining or related to the pictorial information, the narration shouldn't be accompanied by music. In other words when there are three media items (picture, spoken text and music) the picture and the spoken text should be chosen for presentation. The music can be left optional. The theoretical explanation according to Mayer's dual-channel model (Fig. 1) is that the music in the background will slow down or hinder the process of selecting relevant verbal and visual material for further processing in working memory because some portion of the recipient's attention will be directed to listening to the music. Moreover background music will be processed by the auditory channel which, in the described situation, needs its main capacity for complete processing of the narration.

Mayer did not conduct experiments testing whether the presence of background music will lead to worse performance results when the combination of pictorial information and on-screen text is employed. We can speculate that in this case background music will also be considered extraneous information as it does not directly participate in the formation of the image base or the text base. And yet there might be users who would prefer to read a text and listen to music. There might be also theoretical and empirical evidence for music enhancing the cognitive processes of reading a text or watching a picture. Modality principle

According to Mayer in a multimedia presentation words should be presented as auditory narration rather than as visual on-screen text. Tabbers (2002) did a series of experiments which led to the refinement of Mayer's modality principle. Tabbers proved empirically that Mayer's modality principle cannot be generalized for all cases, but is only applicable for system-paced instructions, that is when the user has predetermined limited time of access to the information, respectively limited time to process it. When the user is left to pace his/her own time to study the textual and pictorial information, the results are in favour of on-screen presentations. These findings can be translated in the following presentation rules:

If for a certain topic the available media types are pictorial information, written and spoken text, where the written and spoken text present identical information, the pictorial information should be combined with either the written or spoken text but not both. Along with the specifically chosen form of verbal information presentation the user should always be provided with the possibility to customize his/her presentation with the available alternative option. Furthermore the rule needs to define the conditions under which decisions for the choice of written or spoken text will be taken. We can argue that for users with more informative tasks (those who wish to spend less time on a presentation and yet grasp as much information as possible) the optimal presentation is that of picture and spoken text. For users who seek deeper comprehension of the information and are also willing to spend more time in studying the presentation in detail the combination picture with written text is optimal. This assumption is based on the split-attention theory described in the cognitive load literature.. The above assumption is not empirically proven..

3. More concrete rules for combining different multimedia types

As an input I am considering the following categories of media information: textual information, visual information and auditory information. The textual information is on-screen text. The visual information is picture or animation. The auditory information can be narrative or music. I also assume that for a certain query the retrieved media items will be thematically related.

Following I consider all the possible combinations of retrieved media items and how they should be combined in accordance with the theoretical considerations presented earlier in this writing.

4 media items

text, picture (animation)¹, auditory narrative and music

¹ Whether either picture or animation or both will be chosen for a final presentation is a consideration out of the scope of this work

If all possible media items are retrieved then the input for a generation process will comprise of text, picture (animation), auditory narrative and music. For an optimal combination of those elements additional information from the user model is needed. The user model needs to specify whether the presentation is intended for fast and rather informative visiting or for comprehensive study. In the former case pictorial information is best combined with auditory narration (see previous section, Tabbers) where the auditory narration does not include background music.

For each instance of generated presentation those media items which were not chosen in the generation process should be made available for the user in case he/she is not satisfied with the generated version and would prefer to customize it. In the case of the presentation described above the choice of on-screen text instead of auditory narrative should also be available as well as the possibility of adding background music.

If the user model identifies a user who is willing to spend the necessary time studying more in depth the presented information then an optimal combination of the available media items will be a picture accompanied by on-screen text. If the user model contains information about preferences regarding background music then this preferences should determine the presence or absence of background music. If this information is not available then background music should be excluded with the possibility that the user adds it on demand.

3 media items

picture, auditory narrative, music

If the input media items are picture, auditory narrative and background music and the user model specifies if background music is user preference then in accordance with this preference a combination of the three items or a combination of picture and auditory narrative can be generated. In the latter case a possibility for “switching on” music can also be provided in case the user changes his/her mind in the process of work.

If the user model does not provide the input necessary for taking a decision about the third media item (music) then a presentation of picture and auditory narrative should be generated where the possibility of including music by the user is also provided.

picture, text, music

This situation is resolved in the same way as the one described above.

picture, text, auditory narrative

Applying the modality principle of Mayer and Tabbers for the proper combination of these three media items information from the user model is needed. For users with informative tasks the optimal combination is picture and auditory narrative for those with more comprehension oriented goals picture accompanied by text will lead to better cognitive processing. And again the media element not explicitly revealed as a result of the generation process should be available to the user for adding.

If the user model does not contain the information that determines the proper combinations then we can speculate that the picture and the auditory information is an optimal initial generation as a first interaction with any information is mainly introductory. In this case the possibility of substituting or adding auditory information to text by the user is of great importance.

text, auditory narrative, music

The proper combination of these 3 media items depends entirely on the user model. Some users prefer to read and listen to music, some prefer to read in silence some prefer to listen to the information with some background music. Though I don't have any theoretical or empirical evidence to cite I think that seldom one would prefer to read the text he/she is listening to because the speed of reading differs for different people and it rarely coincides with the speed the text is read through auditory narrative.

If there are no indications from the user model on how these media items should be combined then text should be presented together with the possibility to change with or add to audio as well as an option of adding music. This seems to be the only reasonable combination as the alternative is presenting an empty screen with running audio and a possibility to change with or add text and/or music.

2 media items

The possible combinations of two media items are the following:

Text and picture

No external consideration should be taken into account for this combination.

Text and auditory narrative

As mentioned above we cannot consider this combination optimal. An optimal presentation in this case will be of the text item alone with a possibility for the user to change it into auditory or have them both run simultaneously.

Text and music

It depends on the user model whether these two media items can be combined. If the user model does not specify the optimal combination then the text should be presented alone with a possibility to add the music.

Picture and narration

This is an optimal presentation.

Picture and music

The same combinations as with text and music.

Narrative and music

The combination is rather problematic because if only these two media items are available for presentation there will be nothing to appear on the screen (neither text nor picture)

There are some further considerations that can be taken into account, for example the synchronization of the different media items. Music shouldn't play longer than the time allocated for presenting all the available pictorial information on the presented topic.

Another interesting theme for investigation is the possible contact points between the different media items, for example two pictures changing at a certain point in the narrative to illustrate adequately the information the narrative is presenting. I consider, for the moment, such topics out of the scope of the present work

References

Baddeley, A.D. (1992). Working Memory. *Science*, 255, 556-559

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, 293-332.

Mayer, R.E. (1999). Research-based principles for the design of instructional messages: The case of multimedia explanations. *Document design*, 1(1), 7-20.

Paivio, A. (1986). Mental Representations: A dual-coding approach. New York: Oxford University Press

Dwyer, F.M. (1978). Strategies for improving visual learning. State College, PA: Learning Services.

Tabbers, H.K., (2002). The modality of text in multimedia instructions. Thesis