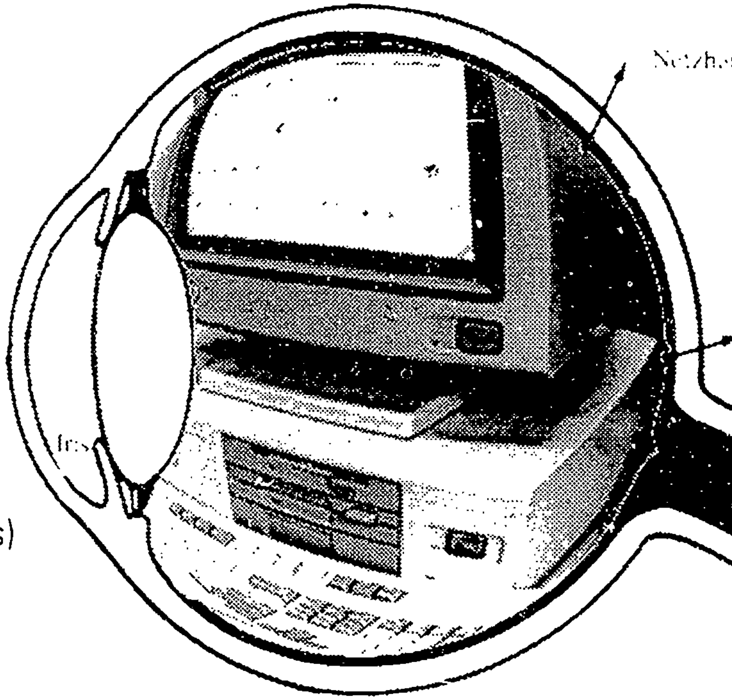
Principles and tools for instructional visualisation



Jan Schoenmaker
Ivan Stanchev (eds)

faculty of
EDUCATIONAL SCIENCE AND TECHNOLOGY

Andersen Consulting - ECC
The Netherlands

University of Twente

ANDERSEN CONSULTING
ARTHUR ANDERSEN & CO. SC

Functions and Design of Video Components in Multi-Media Applications: a Review

Pløn W. Verhagen

Introduction

Designers of interactive multi-media applications face the decision when and how to be presented information has to take the form of text, graphics, animation, video stills, moving video, sound, or any combination of the former. This contribution focusses on design decisions with respect to the use of live or animated video segments with or without sound. First, functions and utilization patterns of video as a component of multimedia systems are presented. Next, research on audiovisual design is reviewed with respect to camera factors, audio factors, optical effect and special effects, pacing and rhythm, animation, picture complexity, and using two channels (visual and audio).

Video functions

Video functions may be classified according to content-driven or logistic criteria which may both be used to support decisions to select video as a component of an instructional application. Figure 1 provides a list that is based on Verhagen (1992, 1993). The functions are:

Content-related use of photographic pictures or moving video (showing people and objects). This function concerns the message-specific need for a certain mode of presentation (if I require a visual, audio will not do; and if I have to show motion, video has to be preferred over still pictures). In certain fields, such as medicin, biology, and geography, there are many content elements for which it is self-evident that they need these kinds of visuals. It has to be noted, however, that in many cases message characteristics are not unequivocally prescriptive with respect to communication mode. Visual referents in memory may for instance inhibit the need for pictures, if they sufficiently support mental imagery in the context of verbal messages. Figure 1 shows that content relatedness plays a role in several of the other listed video functions.

Depicting the invisible or the non-accessible. This concerns presentation options that cannot exist without certain media techniques. Examples are: slow motion, time- lapse photography, photography with invisible beams (like Röntgen or infrared photography), use of telescopes and microscopes, and animation. Here the "human size" with respect to perception in space and time is an important factor. Some experiences cannot happen if events could not be speeded up, slowed down, scaled up or scaled down to manageable proportions by use of media.

Logistic use. Straightforward applications concern uses of media to solve problems of place, time, and identical repetition of messages. (In most cases, it is more feasible to show a slide of the Vesuvius than to visit this volcano). And, substantial effort to develop an audiovisual programme which contains a refined discourse about a certain subject is a one-time investment. (The programme can be repeated infinitely with constant quality.) Logistic arguments apply also in the cases of several other functions, as is represented in Figure 1.

Video functions:	Purpose: Content-related	Logistic
Content-related use of photographic pictures or moving video	•	
Depicting the invisible or the non-accessible.	•	
Logistic use.		•
Cognitive help (visualizations as tools for thinking).	•	
Modelling of psychomotor skills.	•	•
Video feedback	•	•
Providing observation materials.	•	•
Psychologically realistic simulation.		•
Video to present gifted teachers.		•
Video design for emotional involvement.	•	

Figure 1: Videofunctions in relation to content or logistics

Cognitive help (visualizations as tools for thinking). Visualizations may offer new perspectives on knowledge or extend existing knowledge in the way that Bohr's Atom Model shaped thinking about how an atom is conceived. Bohr's model provided a way for insightful communication about the elementary decomposition of matter: a message for which visual media characteristics could be exploited succesfully. Similar reasoning applies, for instance, to the insightful animations (with narration) that demonstrate the essence of the Lorenz transformation or Einstein's Special Theory of Relativity. In general, schematizing is the keyword here.

Modelling of psychomotor skills. Video is in many cases an appropriate means to demonstrate motor behavior.

Video feedback. Video may be used to record student behaviour for feedback purposes. This may support the evaluation of psychomotor performances, feedback during microteaching sessions or interview training, and other.

Providing observation materials. Video registrations (of group discussions, of children playing, of animal behaviour, etc.) may provide opportunities to train observation skills and to learn to analyse behavioural patterns.

Psychologically realistic simulation. Social interaction may be simulated by having human models on -- for instance -- videodisc who react to decisions of students by talking directly to the students from the video screen. Well-known examples are video-based management games.

Video to present gifted teachers. In fact a variant of logistic use, this video function provides the possibility to repeatedly enjoy the presentation of outstanding teachers who demonstrate their skill in handling subject matter.

Video design for emotional involvement. In this case, aesthetic value, emotional value, and use of drama are at stake. Factors regarding pleasure, goodness, beauty, interest, and complexity can be identified as components of aesthetic value. These factors play a role in determining the attention of learners and their perception of the relevance of a learning task which is a component of their motivation. The depicted objects may have emotional qualities of their own. Wounded victims of war or hungry people of Africa may evoke emotional responses that may have an impact for certain types of learning. Responses of anger, joy, curiousity, desire for justice, and other, may also be stimulated by appropriate use of drama. The filmic form of video offers excellent opportunities to exploit these possibilities. Recorded material has thereby the advantage of the carefully planned impact that results from skillful design and production.

Basic approaches for utilization

The above-mentioned video functions represent one set of variables to consider when designing practical application. Next to this, several basic utilization patterns may guide the designers:

Text primacy (symbolic first). Here, the basic question is: "Can printed material (text and pictures) do the job?" This starting point leads in many cases to cost-effective solutions. In cases where moving video is indispensible, this approach may be maintained by using barcode-controlled videodiscs as an adjunct to the written material. The approach is feasible for good readers with sufficient learning skills concerning the subject matter involved.

Visualization as starting point (iconic first). Here, the basic question is: "Can we show what we mean?" This starting point leads in many cases to a look-and-tell format on the basis of video presentations. It results in concrete presentations that may very well suit rather concrete subject matter (or it may enlighten abstract subject matter by visualization). It is feasible for group instruction as well as for individual study, may help poor readers, and may involve students that are not very motivated to read.

Between-channel redundancy as reinforcement strategy. Using two sensory channels increases the likelihood that messages arrive. This phenomenon may be used by developing educational materials in such a way that sight and sound are simultaneously used to present semantically overlapping information. The simplest (and most redundant) form is that a voice presents texts that can be read from paper or from a computer screen at the same time. Striving for semantically overlapping information leads often to audiovisual presentations in which schematizing is frequently used to support abstract subject matter with appropriate images. Between-channel redundancy as reinforcement strategy is appropriate for complex subject matter and/or poor learners as long as the total amount of information does not exceed the information processing capacities of the learners.

Orchestration for learning (theater of life). This concerns the use of the video function for emotional involvement. Creative communication solutions may attract and maintain attention, establishing learning situations that are entertaining, challenging, or have an emotional impact for other reasons. This technique is appropriate if motivation to learn is considered a problem, for instance in distant learning situations where the delivery of instruction takes place by means of television and has to compete with other television channels.

Audiovisual archive as a starting point (providing a computerized information environment). In this case the database capabilities of multimedia are exploited. This may be done for the purpose of knowledge acquisition in an eductional setting, where personal development is paramount. Free navigation through the information space is here the starting point albeit that the space available may be limited for didactical reasons to prevent novice learners from being overwhelmed by the difficulty level of encountered information or from getting lost in the possibly vast amount of information contained in the system. In all cases, adequate help functions should be available. For experienced users, playing around in the information environment offers ways to extent and deepen knowledge by enriching existing schemata with new facts or new combinations of facts.

An other application concerns problem solving on an individual basis with the multimedia system as the main source of information. In that case two more options are there: (a) the problem-solving activity is guided by the system (it is incorporated into a detailed planned teachin g method), or (b) the system is an open information environment in which the user has to find his or her own way. The level of support by the system may be adjusted to the needs of the user, either automatically (with an on-line help system that detects ineffective conduct of the user), or on request (when the user decides to switch from the browse mode to "challange me", "teach me", or "question me"). Applications may further take the form of interactive encyclopedia, interactive maintenance manuals, etc.

Interactivity as the starting point. This is the case when learning outcomes are desired for which it is considered to be necessary that practice, testing, and feedback are organized in interaction with the computer system. The main instruction functions are: orienting on the learning task (presenting subject matter, presenting content-specific ways of thinking, presenting new methods and techniques to operate on the subject matter, demonstrating working methods and problem-solving approaches), practice (providing opportunities to practice, providing feedback), testing and general feedback. For the organization of multimedia applications to serve these instruction functions, common computer-based instruction formats may be used, such as drill-and-practice, tutorial, and simulation, extended with the presentation richness of multimedia.

The video functions and the basic utilization patterns are not independent of each other. Figure 2 shows how the different functions relate.

Basic utilization patterns	Supported by the following video functions
Text primacy	- Limited supplementary use of content-related video
Visualization as starting point	- Content-related use of video - Depicting the invisible or the non-accessible - Logistic use of video - Cognitive help - Modelling of psychomotor skills - Providing observation materials - Video design for emotional involvement
Between-channel redundancy as reinforcement strategy	- All functions with proper audio layers added
Orchestration for learning	- Video design for emotional involvement
Audiovisual archive as a starting point	- Content-related use of video - Depicting the invisible or the non-accessible - Logistic use of video - Cognitive help - Modelling of psychomotor skills - Providing observation materials - Video to present gifted teachers
Interactivity as a starting point	- Modelling of psychomotor skills - Video feedback - Providing observation materials - Psychologically realistic simulations

Figure 2: Video functions from Figure 1 in relation to basic utilization patterns

The video functions and the basic utilization patterns form a limited set of factors to consider when selecting or developing multimedia applications. For instruction, proper handling of learning objectives, content analysis, and learner characterisitcs require thorough knowledge and skills with respect to the knowledge domain of the subject matter and of methods and theories of instruction, and knowledge and skills with respect to methods for multimedia design and production. These subjects are beyond the scope of the present paper. This contribution focusses on results from audiovisual research that relate to detailed design decisions for developing video presentations. This concerns primarily the utilization patterns "visualization as starting point", "between-channel redundancy as reinforcement strategy", and "orchestration for learning".
Much of the audiovisual research reported is carried out with linear video programmes. The findings are, however, probably to a large extent applicable to components for multimedia solutions.

Results of audiovisual research on presentation variables

Presentation variables are actually production variables that obtain their values during the production of video material. Below the following variables are discussed (derived from a list by Coldevin, 1981, p. 87):
(a) camera factors (angle & shot);
(b) setting;
(c) audio factors;
(d) special effects;
(e) pacing and rhythm.
In addition, a brief passage is devoted to animation. The variables are listed in Figure 3 in which they are positioned relative to the production phase in which the pertinent design decisions are taken.

Camera Factors
As far as research is concerned, the effects of camera factors on learning seem to be limited and results are often conflicting. Dwyer (1978, p. 168), for instance, refers to Roshal (1949) and McCoy (1955) to conclude that it is advisable to show a performance on the screen the way the learner would see it if he were doing the job himself. Coldevin (1981, p. 88), however, describes an experiment by Grant and Merrill (1963) that shows for relatively complex skills (and perhaps for most task-oriented productions) that the viewing angle of the demonstrator should not be used rather than the perspective of the student viewer who is watching the demonstration.

Research by Salomon (1974) seemed to show that certain camera factors -- such as zooming in -- can positively supplant visual information that young viewers must fill in during cuts between shots. It has, however, been shown that this supplanting effect can be demonstrated by other means (Bovy, 1983).

This effect is consistent with the fact that different authors often present different solutions to serve certain communication function by camera factors. Zooming-in for instance, is classified by Morrison (1979, p. 29) as a device for: focusing attention, relating parts to a whole, emphasizing one aspect, and showing spatial relationships. All these function can be accomplished by other means. Morrison himself says that the cut can be used to (re-)focus attention and to relate parts to the whole.

But there seems no reason to exclude the use of cuts for emphasizing and for showing spatial relationships. Moreover, for all these functions, the superimposition of optical markers can also be used (such as arrows, animation, encircling; see for instance Dwyer, 1978, p. 160). Lumsdaine and Sulzer (1951) had earlier shown the effectiveness of devices of this kind for directing perception in films. Boeckmann, Nessmann, and Petermandl (1988) confirmed this effect as part of an experiment in which subjects had to watch a video programme with the task to notice objects and procedures that are forbidden in a professional kitchen for hygienic reasons. In the experimental version of the programme, three things were made to stand out in a long shot by means of

Design decisions in production phase:	Presentation variable	Purpose	Examples
scripting and shooting	camera angle & shot	framing of to be depicted objects	close-up to focus attention, wide-angle shot to establish a visual context for a beginning scene, etc.
	audio (narration, live sound, and music)	supporting the visual layer or (in case of text) mainstream of factual information	narration to explain what the picture shows, live sound to demonstrate how a MIG-welding arc should sound, music to support a story-relevant mood.
scripting and editing (post production)	optical effects/special effects	getting attention, relating scenes, symbolizing transitions in place or time	wiping to connect mutually supportive or opposing visual situations, lap dissolve to move to a scene in the past, fade out/ fade in to suggest the passing of time, etc.
	pacing & rhythm	providing a presentation rate that gets and maintains thougthful attention (by complying with the spectators preferred state of cognitive pace)	shorting time by inserting only action-laden shots, edited for narrative continuity at an attention-demanding pace.
scripting, shooting, editing	animation	attention gaining and visualization	instructional visualizations without distracting details, focussing attention by animated pointing devices

Figure 3: Selected production variables for video materials

superimposed rings (flowers, cigarettes, and domestic animals). The results of that experiment show that: "the most impressive increase in recall was observed in the case of the point concerning flowers, these being difficult to see in the background of the picture (38% of the students who had seen the experimental version mentioned flowers in comparison with none of the students who had seen the original version)" (Boeckmann, Nessmann, & Petermandl, 1988, p. 118).

This supports the notion that camera factors do not have definite meanings that can be used as invariable building blocks in message design. Just like the ideas about editing that are put forward by Reisz and Millar (1981), the use of camera factors should be guided by the nature and order of the message components that are to be conveyed. Studying the effects of single shots seems not to be very useful in that respect. It is more that sequences of shots should establish adequate patterns to cover the scenes to be depicted. This is exactly what media handbooks offer to support the development of solutions for particular design problems. Arijon (1976) for instance presents hundreds of model solutions for a wide range of camera problems, varying from staging straightforward two-person dialogues to staging complicated mass scenes or finding solutions in confined areas such as the interior of an aeroplane. It seems thus warranted to approach the treatment of the camera from the perspective of staging a message as a coherent series of shots. Consistent authorship may thereby lead to a style that establishes clarity and appeals the audience. The following research results are consistent with this view:
Low angle shots can increase the perceived potency of a presenter, but the framing of these shots in the narrative structure has then to support this intention (Coldevin, 1981, p. 88).
In video programmes, interest levels can be increased with medium close-ups or with close-up shots (p. 88). Boeckmann, Nessmann, and Petermandl (1988, p. 110), however, concluded on the basis of their research that recall was not related to type of shot (close-ups, long shots, or intermediate shots).

These results illustrate that the function of shot type is dependent on the circumstances under which it is used. The camera treatment has thus to be regarded as a flexible component of audiovisual message design, whereby research-based rules to predict good results are weak or absent. Good results are dependent on the professionalism of responsible team members in a media project. Rules from practice such as can be found in media handbooks are useful to support design decisions (see, for instance, Arijon, 1976, or Millerson, 1985).

Audio factors
Audio factors are narration or dialogue (in the case of speaking characters), effects (the 'natural' sounds of the setting that is depicted on the screen (effects may be recorded live or be produced artificially), and music.
For instruction, off-screen commentary is often used. According to Dwyer (1978, p. 169), the number of words per minute of film in the commentary has a definite effect on learning. He refers to Zuckermann (1949a) and Jaspen (1950a, 1950b) to conclude that care should be taken not to "pack" the sound track. Schmidt (1974, p. 332) found in his analysis of the design of 20 outstanding instructional films that the majority of these films had an average narration rate of 140 or fewer words per minute. Dwyer (1978) advises that:

> "Too much or too little talking in words per minute of film has been found to detract from the teaching effectiveness of a film. The optimum word rate is about 100 words for each minute of film." (p. 172.)

The normal rate of speech ranges from 150 to 200 words per minute (Rossiter, 1971). This leads to a rule of thumb that it is a normal value for narration to take about two-thirds of the total presentation time of an instructional film.

Schmidt (1974, p. 332) also found that the majority of the outstanding instructional films he analysed: (a) did not use a style of narration that talked down or lectured to the audience; (b) did simplify the message as much as possible; and (c) used the active form of sentence structure. That 'simplification of the message' was a factor most probably related to the fact that the analysed films were not designed for specific audiences. Still, the message of Schmidt's observations is that narration should take the audience seriously, present in a clear, concise, and unambigious way, and talk directly to the audience if the subject allows it.

About the other factors -- effects and music -- research results are very limited. Fleming and Levie derived one principle with respect to concept learning; that if a concept is basically temporal -- such as rhythm or time or sequence or frequency, or like poetry or music or speech -- then audition is appropriate (Fleming & Levie, 1978, p. 48). Research about sound effects was not mentioned further by the authors reviewed. The informational and emotional nature of this kind of sound seems to be considered as self-evident. Research about music in instructional audiovisual programmes is scarce. When Coldevin (1981) prepared his review, he could refer to only one study (by Baggaley, Ferguson & Brooks, 1980). In that study different types of music for opening or closing a programme were subject to experimentation with unconclusive results due to methodological problems (Coldevin, 1981, p. 89). With respect to music, Jaspers (1991a) reinforced the conclusions made by Zuckermann (1949b) that music has informational, emotional, and conceptual/integrative functions. Leit-motivs and other musical effects may help to structure a presentation and inform the listeners about what is going to happen. Tempo, modality, and rhythm may be used to evoke emotional responses. According to Jaspers, proper use of music may support instructional messages but -- unless the topic is about music -- there is no need to use it "when the purpose is transmission of referential information and cognitive information processing" (Jaspers, 1991a, p. 50). Music may, however, facilitate learning by giving students occasions for identifications and empathy. However, as Jaspers puts it: "well-chosen music does not guarantee to create the supposed effect, but it is highly probable that ill-chosen music will do much harm to the learning process" (Jaspers, 1991a, p. 50).

Given the limited results from research, the main conclusion has to be that the treatment of sound effects and music has to lean on rules from practice. Insp'ring ideas can, for instance, be found in Millerson (1985, p. 353-362).

Optical effects/special effects
The review by Dwyer (1978) as well as more recent research by Boeckmann, Nessmann, and Petermandl (1988) show that optical effects such as fades, wipes, and dissolves do not increase the instructional effectivity of films or video programmes compared to straight cuts to connect subsequent shots. Ginsburg, Bartels, Kleingunther, and Droege (1988) found that with special effects, the situation may even be worse. They carried out experiments with a programme that contained highly abstract visual effects and found that these effects diminish recall. Most probably, the observation of Boeckmann, Nessmann, and Petermandl (1988) applies that formal elements -- such as a fades or wipes -- that do not relate to the programme content have no influence on recall (which according to Ginsburg et. al. may even be negative in the case of special effects). Boeckmann et. al. demonstrated that formal elements that played an explicit role to define a message proved to influence recall. The conclusion seems to be that deliberate content-related use of

optical effects or special effects can be feasible, depending on their functionality that affects the quality of a message. As with the use of the other presentation variables, the message design has to be content driven. This is, however, not to deny that form influences content. The eventual meaning of a message is determined by the combination of the depicted and the manner of depiction. It may be stressed that for instructional purposes, form and content should be balanced with the intention to reinforce the substance of the instructional message.

Pacing & rhythm
A question is, in what way does the pace of presentation in films or video programmes have to be controlled to account for cognitive pace (the rate of mental processing of incoming data)? One principle is that: "the rate of development of a film should be slow enough to permit learners to grasp the material as it is shown (Jaspen, 1950a; Ash & Jaspen, 1953)" (Dwyer, 1978, p. 168).

According to Dwyer this means that the rate must be slow rather than fast. The outstanding instructional films that were analysed by Schmidt (1974):

> "... had a rate of development that was slow enough for the viewer to grasp the material as it was shown" (p. 332). "... had a slowing of the rate of development at points at which it is necessary for the viewer to change attention from one source of information to another" (p. 332). "... had an average shot length of about 10 seconds" (p. 334).

The 10 seconds are of the order of magnitude of the time that Simon (1974) has put forward as the time it takes to fix a chunk of information in long-term memory. The same 10 seconds seem also to be candidate for a characteristic parameter in a study by Leahy (1983). Leahy studied the relationship between visual realism and information processing time in viewing projected visuals. His results, with respect to manipulating projection time to promote more efficient visual comprehension, did not yield clear advice. He got, however, the impression that the fact that he did find significant differences between projection times of 5 and 10 seconds and between 5 and 15 seconds, but not between 10 and 15 seconds, that after 10 seconds a ceiling affect might have occurred. Whether this means in general that a mean cutting rate of six shots per minute yields an adequate pace of an instructional film can, however, not be concluded.
Coldevin (1981) referred to a study by Schlater (1970) that showed that cutting rate should not be too slow:

> "Schlater (1970) produced five videotapes in which the visual changed ranged from one every 30 seconds in the first treatment to nine per 30-second interval in treatment five. Three types of learning tests were administered: pictorial video (visual sketches), verbal video (verbal descriptions of pictorial information) and audio channel information. Testing results showed that verbal video comprehension increased directly with the increase in the rate of visual changes (levelling off at seven visuals per 30-second interval), but the opposite effect was noted in pictorial video recall. As might be expected, no significant differences were found between the treatments on recall of audio channel information." (Coldevin, 1981, p. 91)

It thus seems that under most circumstances a cutting rate that corresponds with a mean shot length of about 4.5 seconds may be feasable.
A further study by Vincent, Ash, and Greenhill (1949) studied the relationship of length and fact frequency to effectiveness of instructional films. "Fact frequency" may thereby be conceived as a factor that is directly related to the cognitive pace mentioned earlier. The procedure in the study was as follows:

"Four experimental film versions dealing with the causes and manifestations of the weather were made up from a series of Navy training films on aerology. The Long Heavy version ran 29 minutes and contained 224 facts; the Long Light version also ran 29 minutes but contained 112 facts. The Short Heavy version ran 14 minutes and contained 112 facts; the Short Light version ran 14 minutes but contained 56 facts. The total number of words in each pair of equal length kept constant by the use of repetitions, prefatory statements, and other 'filler' material which did not add new facts. The four experimental versions were shown to four groups in each of three different populations, High School students (12th grade), Air Force basic trainees, and College students. In each population a fifth control group did not see the film. All groups took the same 136 item multiple-choice test. The High School and Air Force groups took the test again after delays of four weeks and seven weeks respectively." (Vincent, et al., 1949, p. 1.)

The pace in the heavy versions was thus about one retainable fact per 7.5 seconds. In the light versions about one retainable fact per 15 seconds was presented. The reported narration rate was about 124 syllables per minute. The results showed that significant learning occured:

"Every group that saw experimental film earned a substantially higher score than the control group which did not see a film. ... For the High School sample the Short Heavy version seemed to be most effective, for the Air Force and College samples the Long Light version seemed to be most effective. At the end of the delayed recall period all differences among the versions were much smaller than they had been on the immediate retention test, and most of them were not significant." (Vincent, et al., 1949, p. 1)

Test scores on immediate recall ranged from slightly more than 40% to about 50%, in the experimental groups; delayed test scores were roughly speaking 20% lower. Differences between groups were of the order of magnitude of 0 to 5%. Vincent, et al. point at the fact that the films were rather difficult for the populations used. Their overall conclusion is that:

"It seems clear from the data that packing more and more information into a film yields only very slight increments in total measured learning (underlining omitted, pwv). In no case did the Long Heavy film group learn anything approaching twice as much as the Short Heavy or Long Light groups, nor did the latter learn twice as much as the Short Light group." (Vincent, et al., 1949, p. 2)

With respect to the audio channel, the narration rate that was discussed under "audio factors" also plays a role in pacing a programme. The narration may "drive" the processing of visuals more when the two information channels are sufficiently correspondent (Grimes, 1990, see also the passage about two channels below). Clear research-based rules on how to proceed when deciding on cutting rate and narration-related pacing, however, do not exist.

Animation

Animation is traditionally a representation technique from the film and video world. It varies from simple stop-motion formats in which message components are stepwise developed (in instructional situations usually synchronous to narration) to detailed cartoons in full motion. The technical developments of computers have led to beginning possibilities to integrate animation into computer-based instruction.

Rieber (1990) posits that in computer-based instruction animation has been used to fulfill or assist one of three functions: attention-gaining, presentation, and practice. Rieber sees reason to advise that animation should be incorporated only when its attributes are congruent to the learning task. These attributes are: visualization, motion, and trajectory (= the path of travel of an animated object).

In the audiovisual world, the situation is not much better as far as research is concerned. The usefulness of animation techniques seems, however, to be accepted as rather self-evident. Animation is considered to be a common presentation option that can be put into operation whenever its application seems feasable. The functions mentioned by Rieber are then also at stake, as are even simpler applications such as providing cueing devices like superimposed boxes and arrows.

An important feature of animation is that animated visualizations can be strictly to the point, without any distracting detail. Wember (1976), for instance, did an experiment with a video report about the political and societal conflicts in Northern Ireland. In this report one sequence was about the historical causes of these conflicts. The images that accompanied this sequence were originally scenes of a town (street scenes, graffity on walls, etc.). Wember replaced these images by an animated scheme of the historical development (on the basis of a highly stylized map of the British isles) that proceeded synchronous to the narration. In this way he was able to demonstrate a dramatic gain in recall and comprehension scores with respect to the intended message (from an initial score of 30% correct with the original version to a score of 80% correct with the version in which the live scenes were replaced by the animation). Wember suggested that the effect of the animation was especially convincing because of the fact that the supported message was abstract. It is generally difficult to find realistic images that support abstract messages that are contained in the commentary. Animated schematic drawings may in that case offer the solution and add to the impact of the presentation.

Aspects of information load

Next to presentation variables, the communicative features of video presentations depend on the cognitive load they cause in the mind of the viewers. Picture complexity should not be beyond what a learner is able to understand and the audio and video channel should not be thus information loaded that they compete for the attention of their audience. In this section, relevant research is reviewed leading to a list of considerations that designers should keep in mind when working on video segments (Figure 4).

Picture complexity

"Picture complexity" refers to the syntactic, semantic and pragmatic difficulty of comprehending a picture. Other factors also can determine complexity:

- Density of information in the sense of both the number of picture elements or objects and of the degree of detail with which each element is depicted (Peeck, 1987, p. 137).
- Syntactic cues such as overlap of objects to augment the depiction of depth.
- Semantic clarity, which refers to the ease with which picture elements can be called by name.
- Pragmatic clarity, which concerns the kind and amount of existing knowledge that is needed to interpret the picture.

Each of these factors have been subject to research with respect to a diversity of variables that can be identified to play a role. Overviews can be found by Goldsmith (1984, 1987), Peeck (1987) and Leahy (1983). Here, only some general issues are discussed with respect to the use of realistic video images as well as the use of schematic representations.

The effectiveness of realistic pictures for learning has been a matter of debate for a long time:

> "An ongoing controversy exists in the literature regarding the role of realism in visualized instructional materials. A number of theorists, writing during the 1940s and early 1950s, suggested the existence of a realism continuum along which all materials could be scaled in relation to their instructional effectiveness. ...
> However, more recent researchers ... have disagreed with these theories on the grounds that more realistic materials will tend to present too many irrelevant cues which will interfere with learning. Miller (1957) suggests that too many irrelevant cues may actually compete with one another for attention and subsequent information storage. ... for some educational objectives, simplified or less realistic visual instructional materials are more effective in falicitating student achievement than more complex photographic or detailed drawings of the same object. Dwyer found strong evidence to indicate that the most realistic visual illustrations were not necessarily the most effective instructionally." (Parkhurst, 1975, p. 175-176)

Heuvelman (1989, p. 38-40) discussed the results of several studies that suggest that for instruction, schematic visualizations yield better learning results and are thus preferred. Dwyer (1972) provided arguments why realistic images may hinder learning in the case of externally paced presentations (as in video programmes):

> "The initial impact of the excessive realistic detail may be sufficiently strong to detract attention from relevant and important learning cues. ... The process of indentification and discrimination is time consuming; the more intricate the visual stimuli, the longer it takes for the student to identify and absorb the intended information." (Dwyer, 1972, p. 90-91.)

The research of Dwyer and others focused on the dimension of realism-schematism. This focus may, however, distract from the real instructional issue. The question is whether the instructional goals need the presentation of concrete referents (to teach concrete concepts, to demonstrate psychomotor procedures, to embed instruction in a familiar context; or any other application for which it is necessary to provide real-world schemata that require "gestalts" with proper fidelity) or whether schematic visualizations are better to articulate the instructionally relevant features (visualizing structures, relations, and principles).

In the case that real-world schemata are indicated, results other than those in favour of schematic visuals may be found. Indeed, Leahy, in his research with fixed-pace presentation of visuals, found that complex high-realism visuals promoted a significant increase in visual comprehension over low-realism visuals Leahy used colored visuals of high complexity on two levels of realism that were of relevant instructional content to social studies, with emphasis on object-interrelationships and conceptual content, rather than mere object depiction. Subjects were public high school sophomores. The properties of the high-realism visuals that seem to be responsible for his results may, according to Leahy, be "a more meaningful, coherent, or experientially recognizable pattern of object-interrelationships" in these pictures (Leahy, 1983, p. 107).
With respect to picture complexity, Leahy (1983, p. 111) listed as one of the implications of his study that high-realism visuals are apparently more efficient with respect to the recall and comprehension of presence and appearance of information and of the interrelationship of objects.
Next to this, another argument for some degree of complexity is that complexity invokes interest, as is apparent from one of the research-based principles of Fleming and Levie: "attention is drawn and held by complexity, providing the complexity does not exceed the perceivers' cognitive capacities" (Fleming & Levie, 1978, p. 22).

A last point is that picture quality has to be such that the images can be appreciated without technical imperfections. Fleming and Levie formulated one of their principles for instructional message design: "the better an object or event is perceived, ... , the more feasible and reliable will be further cognitive processes, e.g., memory, concept formation, problem solving, creativity, and attitude change" (Fleming & Levie, 1978, p. 88). The use of professional video formats and professional camera treatment and editing facilities seem thus to be indicated in order to produce effective video materials.

Using two channels

In audiovisual presentations, the visual channel and the audio channel are used simultaneously to convey messages. Several researchers have tried to find how these two channels interact. Jaspers (1991b) reviewed research regarding the relationship between sound and image. His findings, supplemented with findings of Dwyer (1978) and Kozma (1991), are that:

"Visual images are reflections of physical presence" (Arnheim, 1986), presenting "concreteness" (Mast, 1977) that is represented by the surface of the depicted objects and events. Sounds are, in most cases, attributes of visual objects (Metz, 1975), with the exception of sound objects such as a hiss, thunder, music (Jaspers, 1991b, p. 164). But, according to Metz, voices may also be aural objects as far as they have meaning of their own, apart from the words spoken (recognizing an off-screen voice may be important to value the meaning of its speech). For instructional applications this is, however, a minor point in most cases or even not relevant.

Language communicates thoughts (Arnheim, 1986); it presents abstractions (Mast, 1977). This counts whether or not language is presented visually (in print or on a video screen) or as an audible sound.

When used in combination, most researchers conclude that an audio-visual combination can be more effective than audio alone or video alone (Hoban & Van Ormer, 1950, reprinted in 1970; Coldevin, 1981; Nugent, 1982). Kozma (1991, p. 191) reviewed more than ten studies in which separate audio and visual presentations of a video programme were compared and found that: "in none of these studies did the combination of audio and visual information result in lower recall than recall from either source alone. In most of these studies, the combined use of visual and auditory symbol systems resulted in more recall than visual-only and audio-only presentations." Although it is thus often the case, it not always evident, that audio plus vision is better (Hsia, 1968). The effects of the combination appears to be relative to factors of learner and material (Hoban & Van Ormer, 1950; Johnston, 1987; Barton, & Dwyer, 1987).

Often, dominance of vision over audio is reported. "The overall influence of the motion picture is thought to be primarily in the picture and secondarily in the accompanying language" (Dwyer, 1978, p. 172). Visually presented information would be retained better than aural information. But, for instance, Hsia (1968) concluded that "the question of supremacy of the auditory or visual channel is controversal. Some studies indicate that seeing is better than hearing, others show the reverse" (Jaspers, 1991b, p. 164-165). These results seem to support one of the principles for improving visual learning that was proposed by Dwyer:

> "Both audio and visual elements of films are effective channels of communication. Neither channel is consistently better than the other. Each channel is uniquely capable of conveying certain types of information and the two should be properly integrated." (Dwyer, 1978, p. 172.)

The results also agree with one of the research-based principles of Fleming and Levie:

> "More learning can occur where information is received concurrently in two modalities, e.g., vision and audition or vision and touch, than where received in only one modality." (Fleming & Levie, 1978, p. 107)

In instructional audiovisual programmes, sounds other than language or music are in most cases attributes of visual objects in the sense of Metz' (1975) assertion. (Instructionally relevant sound objects other than music do, however, exist. Examples are ticks that increase in speed to inform learners about the approach of the end of a period for a timed answer, and the different warning signals and other tones that are often used to give feedback on learner actions during learner-computer interaction. For audiovisual presentations, sound applications of these types are not very often used.)

With sounds as attributes of visual images, the relationship between the visual images and language becomes the central issue for instructional audio-visual message design, in particular if the language consists mainly of off-screen narration, which is the case with the often used look-and-tell format. It has, however, to be noted that Boeckmann et. al. (1988, p. 110) found that in their experiments no influence on recall could be established in connection with the fact "whether the main source of sound was shown or not (e.g. whether the commentator was visible or a voice-over was used)". This suggests that conclusions with respect to off-screen narration may apply to presentations by a studio presenter as well.

With respect to off-screen narration two variables are at stake: (a) the correspondence of the visually presented information and the commentary; and (b) the use of narration-supporting superimposed text on screen. Both are treated below.

Correspondance of video images and audio narration

The relationship between the information in the two channels determines whether the proper integration occurs as Dwyer has suggested. Some researchers have put forward the idea that "redundancy" and in particular "between-channel redundancy" plays a role. Jaspers (1991b, p. 168) describes that redundancy is formally defined as one minus the ratio of actual-to-maximum information transmission, but that the common use of the term is much more informal and refers to superfluous, extra, or repetitive information. Hsia (1971; 1977) has opted for optimal between-channel redundancy which he sees as the key to better communication. Between-channel redundancy in audiovisual communication occurs if information that is presented in in the visual channel matches the information that is at the same time presented in the audio channel. A methodological problem appears to be that it is not an easy matter to decide on the extent to which certain information can be conceived to be redundant. Pictures have, for instance, many cues that may be used by learners to connect new information to elements of their existing cognitive structure, while with audible text it is assumed that to be redundant to the picture only one perspective with respect to that picture is offered. This concerns the problem of the limited possibilities of translating information that is represented in one communication mode (for instance text) into a representation in another mode (for instance a visualization). In the past decades, a body of research and theory has developed with respect to sign systems (often also referred as symbol systems, where the term "symbol" is then used as a synonym of "sign"). Gross (1974, p. 60) speaks of primary modes of symbolic behaviour and distinguishes: (a) the linguistic, in which oral or written language is carrying the messages; (b) the social-gestural, in which facial expressions and body language communicate messages between people; (c) the iconic, in which communication takes the form of visualizations; (d) the logico-mathematical, in which the symbolic language of mathematics constitutes an independent way of communication; and (e) the musical, in which people express themselves with musical sounds. They are: "identified with (a) a range of objects and events or field of reference, (b) a distinctive memory-storage capacity, (c) a set of operations and transformations, and (d) specific principles of ordering which govern the formulation and communication of meaning, and (e) nontranslatability into other modes" (Gross, 1974, p. 61). With the last point, Gross means that information which is coded within one mode will not be capable of being fully recoded in terms of another. Nontranslatibility is thus a characteristic of a communication mode which indeed limits the possibilities for between-channel redundancy. It is therefore better to speak of more or less correspondence between pictures and sound, whereby strong correspondence resembles redundancy but is, in fact, not the same. This is here thought to be one of the reasons why results of experimental research fail to demonstrate that striving for redundancy yields results that contradict predictions on the basis of cue-summation theory (see for instance Severin, 1967). Cue-summation theory says that adding non-redundant cues in different channels enriches a message, thus offering more opportunities for learning (as long as the information processing system of the learner will not be overloaded), and redundant cues do not (because they add no new information). If after all alleged redundant information appears not to be as redundant as was assumed, improvement of learning that is reported on the basis of redundancy may very well be the consequence of cue-summation. A second reason why redundant cues may support learning despite of the predictions of cue-summation theory is, that in this theory the didactically beneficial effects of repetition are not taken into account. In practice, most probably, both reasons play a combined role.

Grimes (1990) used a study of Drew and Grimes (1987) as starting point to study the influence of channel correspondence on recall of auditory or visually presented information.

> "Drew and Grimes found that in the optimum viewing condition -- high channel correspondence -- audio recall and story understanding were the highest, and video recall was the lowest. In the least optimum viewing condition -- no channel correspondence -- video recall was the highest, audio recall was the lowest. and story understanding was the worst." (Grimes, 1990, p. 16)

It has to be noted that Grimes' research concerned television news, a format in which the main stream of information is traditionnally contained in the audio channel. Grimes reasoned that the results of Drew and Grimes are a consequence of limited memory and/or attentional capacities:

> "Channel correspondence probably affects the way in which attention is divided between the auditory and visual channels. Because attention is a limited commodity (Kahneman, 1973), more attentional capacity is probably required to encode and integrate in memory the two information channels when those two channels convey slightly to moderately different messages. With enough difference in messages (e.g., a medium-correspondence condition), attention's capacity threshold will probably be approached, maybe even exceeded, and viewers may try to concentrate on the auditory channel -- the who, what, when, where, why channel -- at the expense of the less specific, and less useful, visual channel. ... ...high channel correspondence should allow the viewer to monitor the auditory channel for the who, what, when, and where of the story and to shadow the visual channel for information that supplements the narration and can be integrated into the information supplied by it. Thus, attention to, and memory for, the visual channel should be greater in this condition than in the medium-correspondence condition.

However, if the conflict between the two information channels is great enough (e.g., the no-correspondence condition), attention to the visual channel will override attention to the auditory channel." (Grimes, 1990, p. 16)

This last idea is, according to Grimes, not only supported by the non-correspondence data from the Drew and Grimes study, but also by the work of Colavita (1974) and Posner, Nissen, and Klein (1976) who report that visual stimuli dominate auditory stimuli when pictures and words have no relationship to one another (Grimes, 1990, p. 17). Grimes carried out an experiment in which he focussed on the locus of attention as it was split between the auditory and the visual information channels. "The purpose was to discover which channel, if either, received less attention due to diminishing channel correspondence" (p.17). He used three versions of four television news stories as stimuli. Grimes operationalized "High Correspondence" by a "tight" picture-word match in the experimental material. When, for instance, the narration talked about a needle with the size of a screwdriver that is inserted into a herniated disk in the spine, the picture showed a needle, the size and shape of a screwdriver, about to be inserted into the spine of a patient. "Medium Correspondence" was operationalized by relating the two information channels thematically but with no further correspondence. The narrated sentence mentioned above (about the needle) was then accompanied by a video scene of a patient being wheeled into surgery.

In the "Non-Correspondence" condition, the visual track consisted of a visual potpourri that did not match the audio. "For instance, as the narrator described how herniated disks are repaired, the viewer saw a barge loaded with garbage anchored in New York Harbor. When the shot changed, it showed celebrants of the most recent Chinese New Year, while the narration continued to discuss surgery" (p. 18). Story narrations were the same across all conditions. Attention to the visual or the auditory channel was measured with a reaction-time technique based on visual or tone probes. Half of the subjects were exposed to seven visual probes (NTSC colour bars during 33 milliseconds) nested within each story, and the other half were exposed to analogous tone probes (with a frequency of 10 kiloHertz). In the technique subjects have to react to the visual or auditory signal each time that it appears by pushing a key as soon as possible. The reaction time is a measure of attention. The longer it takes, the more attention is assumed to be devoted to the channel within which the signal appears (due to the channel capacity that is occupied by the attention). Two recognition posttests were used: (a) a factual memory test with 53 multiple-choice questions; and (b) a 36-item visual memory test on the basis of freeze frames from video material that was used in the stories as well as from "outtakes" (shots not included in the stimulus material). Subjects were 202 undergraduate communication majors who worked individually through the experiment.

The results showed no statistically significant differences between factual memory scores in the high- and medium-correspondence conditions (the scores were 56% and 54% respectively). In the no-correspondence condition, factual scores were much lower (35%), but the auditory probe reaction times were not significantly different from these times in the other conditions. "This suggests that subjects were paying attention to the auditory channel, but that they did not remember much of the factual information contained within" (p. 23). Grimes discusses that this is consistent with the earlier mentioned observations of Colavita (1974) and Posner, et al. (1976) and also of Cohen (1973) who described that visual signals appear to go through fewer extrapolations than auditory signals before they result in memory codes. "Visual messages do not need as much attention to be efficiently encoded, and therefore get priority entry, ..., into the human information-processing system" (Grimes, 1990, p. 23). If unrelated auditory and visual stimuli are competing for simultaneous processing, visual encoding primacy then keeps the factual information from being properly processed.

Visual recognition scores for the high-, medium-, and non-correspondence conditions were 66%, 55% and 64% respectively. The high visual recognition scores for the non-correspondence condition can partly be explained on the basis of the visual encoding primacy, that also caused the low factual memory scores in case of extreme audio-video conflict. In this condition, reaction times to the visual probes were relatively short, suggesting that the visuals were processed only to the point of recognition. The difference between visual recognition scores for the high-, and medium-correspondence conditions is statistically significant. It shows that audio-video correspondence influences the total amount of information that can be processed mentally per unit of time. In the high-correspondence condition longer reaction times to tone probes were positively correlated to visual recognition scores ($r = .387$, $r2 = .150$, $p < .025$):

> "This suggests that as subjects paid more attention to the narration, they were able to process the video better, and so they produced better visual recognition scores. In other words, as subjects paid more attention to what was said, they used their remaining attentional capacity to integrate what they saw with what they heard. This result suggests that when the two information channels are correspondent, the narration 'drives' the processing of the visuals." (Grimes, 1990, p. 24.)

This is seemingly in contradiction to the view of Kozma (1991, p. 192), who tends to conclude that comprehension of video appears to be driven by the processing of visual information:

> "This is apparent from a study by Baggett (1984), who varied the temporal order of audio and visual information within a video presentation on the names and functions of pieces of an assembly kit. In this study, the narration was presented in synchrony or 7, 14, and 21 seconds ahead of or behind the visual presentation. College students performed best on immediate and 7-day delayed tests of recall of the synchronous and 7-second, visual-then-audio presentations. The worst performance was by groups with the audio presented first. This suggests that, in a video presentation. the visual symbol system serves as the primary source of information and that the audio symbol system is used to elaborate it."

The two studies combined. however, need not yield a conflicting perspective. It may very well be that learners take the first few moments after switching to a new shot to process the new visual image to a level of recognition that is sufficient to be open to subsequent elaboration which is then guided by narration. In that case, Kozma's conclusion that comprehension of video appears to be driven by the processing of visual information, has at least partly to be rejected.

The question of excessive information load due to lack of audio-video correspondence appeared also in a study by Wember (1976). Wember analysed the informational structure of filmed background reports about the difficult political situation in Northern Ireland. He found that the filmmakers edited the visual layer of these films in such a way that they are dynamically paced and have an attractive rhythm. This had, for instance, the consequence that static shots were held relatively short (mean projection time: 2 seconds), while at the other end of the scale shots with movement of the camera and movement in front of the camera at the same time ran uninterrupted relatively long (mean projection time: 5.9 seconds). The content of the films did, however, often not motive this editing style. Despite the fact that the filmmakers denied this, it seems that their professional consent that "film is movement" played them tricks.

This is an occasion to stipulate that the media profession is differentiated. Media professionals who are schooled to work in the entertainment industry have, for instance, a quite different view of their field than media professionals who are educated to serve educational communication. The need for differentiated schooling is, however, still recognized by scanty measure. The cooperation between filmschools or schools of cummunication and departments of educational technology is still very limited. This is the more problematic, now that multimedia technologies open a market where the combined skills of both worlds are inevitably needed.

Wember perceived that the main effect of the dynamic pacing is that it attracts attention and prevents viewers from switching to another channel when these films are broadcast on television. But he was also able to demonstrate that the superfluous "Augenkitzel" (tickling the eyes), as he called it, demanded so much from the visual processing capacity of the viewers, that the intended, mostly rather abstract messages in the audio channel were not understood or at least not stored into memory although the viewers reported that they thought that they did. He concluded that the gap between pictures and commentary in the films he analysed in many cases forced the viewers to divide their attention between competing channels, with a consequence that they failed to really take in anything substantial of the presented information. Wember's study confirms that lack of correspondence between the visual channel and the audio channel is detrimental for learning. Unlike the findings of Grimes, in Wember's case the recall of the pictures dropped to a very low level.

Wember's results pertain to attention processes in the visual channel that hindered the appreciation of the information in the audio channel. Another mechanism also exists. Boeckmann et al. (1988) assert that picture continuity takes precedence over a change of subject in the commentary:

> "When there is a change of subject in the commentary, even if this should amount to no more than taking up a new aspect of the same subject (such a change as is indicated in a written text by a new paragraph) and there is no corresponding structural marker (e.g. a cut, a graphic, flashing) the continuity of the information carried by the picture conceals the jump from one idea to the next and leads to reduced recall of the point in the commentary. Take, for example, the commentary that moved from 'the optional school subject' to 'the compulsory subject', while throughout moving pictures of a school class were shown. The analysis of the poor recall revealed that the experimental subjects had largely failed to perceive the transition."
> (Boeckmann et al., 1988, p. 112)

Heuvelman (1989) carried out several experiments with respect to between-channel correspondence and also came to the conclusion that "the relationship between the audio track and the visuals is of paramount importance to information processing by viewers" (p. 167). One of the factors he studied concerned the effects of studio presentation on camera in comparison to realistic or schematic visualizations of abstract subject matter. His results led him to conclude that:

> "With abstract subject matter it is very difficult to match pictures with sound when using realistic visualizations. In that case, studio presentation on camera will lead to better results. ... With an increasing level of abstraction of the text, realistic pictures can only show a limited number of aspects or elements of what a certain idea or concept fully means. At the same time, in the eyes of the beholder, realistic pictures can have many potentional meanings. Selection of the intended meaning, i.e. that intended by the maker of the programme, is not imperative. ... In schematic visualizations, a selection, or abstraction, from reality has already taken place: only elements of interest are represented In that case picture and sound will show a better match and, consequently, processing and subsequent retention of the information will be better."
> (Heuvelman, 1989, p. 167.)

The view of Heuvelman implies that lack of between-channel correspondence does not only result in less learning if the information processing capacity of the learners is overloaded, but also if it leaves room for learners to invest their mental effort to derive non-intended meaning. That the use of schematization may help to solve this problem was also demonstrated by Wember (1976), when he used animation instead of live images to support abstract content that was presented through the audio channel.
In all, enough support exists to believe that proper between-channel correspondence promotes learning and that the amount of between-channel correspondence affects the information load of audiovisual presentations. How to control between-channel correspondence is, however, a complex matter. Factors involved are: (a) the concreteness or "visualizationess" of subject matter; (b) the matter of enrichment versus overload; (c) avoidance of distracting information; (d) amount of cueing of auditory content elements by optical markers or changes of shots; and (e) the mutual tuning of audio and pictures in general, inclusive of the wording and timing of commentary relative to the pictures. For every new product, the designers have to make content-driven local decisions to optimize between-channel correspondence with respect to these factors. It goes beyond the scope of this review, to elaborate design procedures to cope with these factors. A few general elements are incorporated in Figure 4 that serves as a conlusion of this section.

Superimposed text
A particular form of between-channel correspondence is formed by superimposed texts that support narration. Heuvelman (1989) reviewed studies of Coldevin (1975), Nasser and McEwen (1976) and Nugent (1982) that show that the presentation of verbal information through two channels (visual and audio) yields information gain compared to audio only, but that no difference was found relative to the presentation of verbal information through the visual channel only (text on screen). This latter form is, however, rather artificial when the use of television is concerned. Heuvelman (1989) himself carried out a few experiments to find out whether short superimposed texts that support narration help recall of presented information. His earlier mentioned conclusion that with abstract subject matter, studio presentation on camera will lead to better results than the use of realistic visualizations was extended with the conclusion that "adding redundant superimposed titles can also lead to better results, but is of less importance and should be applied with care, taking into account the information density of the 'background' picture" (p. 167).

Van der Heijden (1987) carried out an experiment with three conditions: (a) audio supported by short texts on screen against a neutral (blue) background; (b) audio plus live-action images; (c) audio plus live-action images and supporting captions. In this last condition only main points were supported by individual words or short statements near the bottom of the screen. The audio track was the same in all conditions. The programme took about ten minutes to present and contained information about a certain social service. This subject was rather abstract, the reason why the images in the last two conditions were, in many cases, not a 100% substitute for the screen texts from the first condition (which supported the narrated text in a way that resembles the use of the overhead projecter to present information point by point). Subjects were 61 students from the last year of a lower technical school and 62 students from the last year of a lower level high school. Conditions were about equally distributed over school types and learners. A pretest-treatment-posttest design was used. All tests were multiple choice. The results showed that the second condition (audio plus live-action images) did worse than the other two conditions (audio plus text only and audio with live-action images plus text). The reason for this might have been that the imperfect match of the images with the narration appeared a burden to the information processing abilities of the subjects, yielding a less effective use of the two channels than in the case where text on screen was available.

The work of Heuvelman and of Van der Heijden give reason to expect that superimposed texts that support narration may be beneficial for learning.

Conlusions with respect to information load
Research is not very conclusive with respect to picture complexity. A differentiated view emerges in which complexity seems desirable to invoke interest, indicated if a need for concrete (real-life) referents exist, and to be avoided if the presentation of structures, relations, and principles is the central focus of the intended communication. These conclusions are summarized in Figure 4.
The effects of using two channels and superimposed text appear to be dependent of the communicative intention which controls the kind of visuals used, and the relative importance of (oral) text. A few general conclusions are summarized in Figure 4.

Summary

The video components of multimedia applications may serve many different functions, such as providing content-motivated moving images, supporting mental imagery by visualizations, and stimulating personal (emotional) involvement with objects of study. The role of video will further be determined by the way in which video is positioned in relation to the non-video components of the system. In that respect, video may be chosen as the primary form of presentation around which all the other components are organized. But the use of video components may also be limited to support text-based materials where the use of video seems inevitable. Or as just one of the resources that users may encounter when browsing through a multimedia database. In all these cases, the video components should be designed with suffcient clarity to fulfill their roles in the communication process. In this contribution, research was reviewed with respect to production variables that should be properly controlled to accomplish video segments of sufficient impact. In sum, the results are as follows:

Camera factors
Research with respect to camera factors illustrates that the function of shot type is dependent on the circumstances under which it is used. The camera treatment has thus to be regarded as a flexible component.
Research-based rules with respect to camera factors are very limited. A firm rule from practice is that for instruction the camera treatment should be content driven. Each camera angle ought to be motivated by its function in the presentation.

Source of information load	Positive indications	Negative indications
Complexity through pictorial realism	May promote recall and comprehension (Leahy, 1983); invokes interest (Fleming & Levie, 1978)	To many cues may interfere with learning (Parkhurst 1975)
Reduced complexity through schematization	Simplified material is more effective for student achievement (Parkhurst, 1975); schematic visualizations yield better learning results (Dwyer, 1972)	Simplified visualizations may lack meaningfulness (Leahy, 1983)
Combined use of the visual and the audio channel	The combination can be more effective than the use of only one of the two channels (for instance: Kozma, 1991), specifically if the capabilities of both channels are properly integrated (Dwyer, 1978); when the two information channel are correspondent, narration 'drives' the processing of visuals (Grimes, 1990)	The combination can lead to excessive information load, which is detrimental for learning (for instance: Wember, 1976)
Attention-demandig visuals, while the audio signal carries the main message	Viewers are capable of concentrating on the audio channel and are able to learn from it, if monitoring the visuals does not cause cognitive overload (Grimes, 1990)	Visuals get priorty over audio messages during perceptual processing. In cases of too many information, the audio information may not arrive properly in memory (Grimes, 1990)
Reduced load of two-channel messages by schematization	Schematizing abstract subject matter supports understanding and learning of narration-driven presentations (Wember, 1976; Heuvelman, 1989)	
Lack of visual cues to mark transitions from one subject to another		Visual continuation of a scene may mask substantial changes in the narration leading to less effective learning (Boeckmann et al., 1988)
Focussing effect of superimposed text	Superimposed text may support learning (Heuvelman, 1989; Van der Heijden, 1987)	The use of superimposed text is of less importance than the prevention of distracting realistic visual images (Heuvelman, 1989)

Figure 4: Information load of video materials

Narration
Video commentary should take the audience seriously, be presented in a clear, concise, and unambigious way, and talk directly to the audience if the subject allows it.

Sound effects and music
Findings from research with respect to the use of sound effects and music is very scarce. Here too, the main conclusion has to be that the treatment of these variables has to lean on rules from practice.

Optical effects and special effects
Findings from research with respect to the use of optical effects or special effects seems to show that deliberate content-related use of these effects may augment the quality of a message.

Pacing
Clear research-based rules about how to proceed when deciding on cutting rate and narration-related pacing appear not to exist. As with the other audiovisual design variables, in this case the design decisions are also subjected to professional judgement in each specific case.

Technical qualities
Picture quality has to be such that the images can be appreciated without technical imperfections. These requirement can be met by using professional video formats and professional camera treatment and editing facilities to produce video materials.

Correspondance of video images and audio narration
Enough support exists to believe that proper between-channel correspondence promotes learning and that the amount of between-channel correspondence affects the information load of audiovisual presentations. How to control between-channel correspondence is, however, a complex matter. Factors involved are: (a) the concreteness or "visualizationess" of subject matter; (b) the matter of enrichment versus overload; (c) avoidance of distracting information; (d) amount of cueing of auditory content elements by optical markers or changes of shots; and (e) the mutual tuning of audio and pictures in general, inclusive of the wording and timing of commentary relative to the pictures. In all cases, mental overload should be avoided. If the amount of information becomes more than the learner can handle, visual primacy seems to inhibit learning from the audio channel.

Superimposed text
There is reason to expect that superimposed texts that support narration is beneficial for learning.

Despite the research findings and due to the limited direct applicability of these findings, a conclusion is that the judgement of an experienced producer and other experienced media personnel is indispensible in meeting professional standards. This puts media professionals in an important position because of the fact that media design directs the impact of instructional messages. The control over media variables is essential to optimize instructional communication. This factor counts for economic advantages of media, for the motivational effects of media and for media-specific didactics that extend the options for organizing instructional strategies in the sense of Reigeluth (1983, p. 18). It is hoped that the research reported here as well as subsequent research on (interactive) message design will lead to more systemized knowledge to support design processes.

References

Arijon, D. (1976). Grammar of the film language. London: Focal Press.

Arnheim, R. (1986). The images of pictures and words. Word & Image, 2, 306-310.

Ash, P., & Jaspen, N. (1953). The effects and interactions of rate of development, repetition, participation and room illumination on learning from a rear-projected film. Technical Report, SDC-269-7-39, Port Washington, NY: Special Devices Center, Office of Naval Research.

Baggaley, J., Ferguson, M., & Brooks, P. (1980). Psychology of the TV image. Farnborough: Saxon House.

Baggett, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media associations. Journal of Educational Psychology, 76, 408-417.

Barton, E.A., & Dwyer, F.M. (1987). The effect of audio redundancy on students' ability to profit from printed-verbal/visualized instruction. International Journal of Instructional Media, 14, 93-98.

Boeckmann, K., Nessmann, K., & Petermandl, M. (1988). Effects of formal features in educational video programmes on recall. Journal of Educational Television, 14, (2), 107-122.

Bovy, R.A. (1983, April). Defining the psychologically active features of instructional treatments designed to facilitate cue attendance. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Cohen, G. (1973). How are pictures registered in memory? Quarterly Journal of Experimental Psychology, 25, 557-564.

Cohen, P., Ebeling, B., & Kulik, J. (1981). A meta-analysis of outcome studies of visual based instructi on. Educational Communication and Technology Journal, 29 (1), 26-36.

Colavita, F.B. (1974). Human sensory dominance. Perception & Psychophysics, 1[illegible], 409-412.

Coldevin, G.O. (1975). The differential effects of voice-over, superimposition and combined review treatments as production strategies for ETV programming. In J. Baggaley, G.H. Jamieson & H. Marchant (Eds.), Aspects of Educational Technology VIII (pp. 55-65). Bath: Pitman Publ.

Coldevin, G.O. (1981). Experimental research in television message design: Implications for ETV. Programmed Learning and Educational Technology, 18, 87-99.

Drew, D.G., & Grimes, T. (1987). Audio-visual redundancy and TV news recall. Communication Research, 14 (4), 452-461.

Dwyer, F.M. (1972). A guide for improving visualized instruction. University Park, PA: State College, Pennsylvania State University, Learning Services Division.

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

Fleming, M., & Levie, W.H. (1978). Instructional message design: Principles from the behavioral sciences. Englewood Cliffs, NJ: Educational Technology Publications.

Ginsburg, H.J., Bartels, D., Kleingunther, R., & Droege, L. (1988). Cosmos revisited: Just how effective are special effects for instructional communication? International Journal of Instructional Media, 15 (4), 319-326.

Goldsmith, E. (1984). Research into illustration: An approach and a review. Cambridge, England: Cambridge University Press.

Goldsmith, E. (1987). The analysis of illustration in theory and practice. In H.A. Houghton & D.M. Willows (Eds.), The Psychology of Illustation: Instructional Issues (pp. 53-85). New York: Springer-Verlag.

Grant, T.S., & Merrill, I.R. (1963). Camera placement for recognition of complex behaviours. Television in Health Sciences Education, Chapter 6. Washington, DC: US Office of Education.

Grimes, T. (1990). Audio-video correspondence and its role in attention and memory. Educational Technology, Research & Development, 38 (3), 15-25.

Gross, L. (1974). Modes of communication and the acquisition of symbolic competence. In D.R. Olson (Ed.), Media and symbols, the forms of expression, communication, and education. The Seventy-Third Yearbook of the National Society for the Study of Education (pp. 56-80). Chicago, IL: University of Chicago Press.

Heuvelman, A. (1989). Buiten beeld. (Off screen.) Lisse: Swets & Zeitlinger.

Hoban, F, & Van Ormer, E.B. (1950). Instructional film research, 1918-1950. Fort Washington, NY: U.S. Naval Special Devices Center.

Hoban, F, & Van Ormer, E.B. (1970). Instructional film research, 1918-1950. New York: Arno Press & The New York Times.

Hsia, H.J. (1968). On channel effectiveness. AV Communication Review, 16, 245-267.

Hsia, H.J. (1971). The information processing capacity of modality and channel performance. AV Communication Review, 19 (1), 51-75.

Hsia, H.J. (1977). Redundancy: Is it the lost key to better communication?. AV Communication Review, 25 (1), 63-85.

Jaspen, N. (1950a). Effects on training of experimental film variables, Study I: verbalization, rate of development, nomenclature, errors, "how-it-works", repetition. Technical Report, SDC-269-7-17, Port Washington, NY: Special Devices Center, Office of Naval Research.

Jaspen, N. (1950b). Effects on training of experimental film variables, Study II: verbalization, "how-it-works", nomenclature, audience participation, and succinct treatment. Technical Report, SDC-269-7-11, Port Washington, NY: Special Devices Center, Office of Naval Research.

Jaspers. F. (1991a). Music in audio-visual materials. Journal of Educational Television, 17 (1), 45-52.

Jaspers, F. (1991b). The relationship sound-image. International Journal of Instructional Media, 18 (2), 161-174.

Johnston, J. (1987). Electronic learning: From audiotape to videodisc. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kahneman, D. (1973). Attention and effort. New York: Holt, Rinehart & Winston.

Kozma, R.B. (1991). Learning with media. Review of Educational Research, 61 (2), 179-211.

Leahy, M.D. (1983). The relationship among field dependence/field independence, visual realism, and information processing time in viewing projected visuals. Unpublished doctoral dissertation, University of Connecticut.

Lumsdaine, A.A., & Sulzer. R. (1951). The influence of simple animation techniques on the value of a training film. Washington, DC: US Air Force Human Factors Research Lab.

Mast. G. (1977). Film/cinema/movie, a theory of experience. New York: Harper & Row.

McCoy. E.P. (1955). An application of research findings to training film production. Technical Report, SDC-269-7-44, Port Washington, NY: Special Devices Center, Office of Naval Research.

Metz, C. (1975). Aural objects. In E. Weiss & J. Belton (Eds.), Film Sound, Theory and Practice (pp. 154-162). New York: Columbia University Press.
Miller, N.E. (Ed.) (1957). Graphic communication and the crisis in education. AV Communication Review, 5 (3), 1-120.
Millerson, G. (1985). The technique of television production. London: Focal Press
Morrison, G.R. (1979, April). The design of instructional television programs. Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco.
Nasser, D.L., & McEwen, W.J. (1976). The impact of alternative media channels: Recall and involvement with messages. AV Communication Review, 24 (3), 263-272.
Nugent, G.C. (1982). Pictures, audio, and print: Symbolic representation and effect on learning. Educational Communications and Technology Journal, 30 (3), 163-174.
Nugent, G.C., Tipton, Th.J., & Brooks, D.W. (1980). Use of introductory organizers in television instruction. Journal of Educational Psychology, 72 (4), 445-451.
Parkhurst, P.E. (1975). Generating meaningful hypotheses with aptitude-treatment interactions. AV Communication Review, 23 (2), 171-183.
Peeck, J. (1987). The role of illustrations in processing and remembering illustrated text. In D.M. Willows & H.A. Houghton (Eds.), The Psychology of Illustation: Basic Research (pp. 115-151). New York: Springer-Verlag.
Posner, M.I., Nissen, M.J., & Klein, R.M. (1976). Visual dominance: An information processing account of its origins and significance. Psychological Review, 83 (2), 157-171.
Reigeluth, C.M. (1983). Instruction design: What is it and why is it? In C.M. Reigeluth (Ed.), Instructional-design theories and models: An overview of their current status (pp. 3-36). Hillsdale, NJ: Lawrence Erlbaum Associates.
Reisz, K., & Millar, G. (1981). The technique of film editing. London: Focal Press.
Rieber, L.P. (1990). Animation in computer-based instruction. Educational Technology, Research & Development, 38 (1), 77-86.
Roshal, S.M. (1949). Effects of learner representation in film-mediated perceptual motor learning. Technical Report, SDC-269-7-5, Port Washington, NY: Special Devices Center, Office of Naval Research.
Rossiter, C.M., Jr. (1971). Rate-of-presentation effects on recall of facts and of ideas and on generation of inferences. AV Communication Review, 19, 313-324.
Salomon, G. (1974). Internalization of filmic schematic operation in interaction with learners' aptitudes. Journal of Educational Psychology, 66, 499-511.
Schlater, R. (1970). Effect of speed of presentation on recall of television messages. Journal of Broadcasting, 14, 207-214.
Schmidt, W.D. (1974). Analyzing the design of outstanding instructional films. International Journal of Instructional Media, 1 (4), 327-336.
Severin, W. (1967). Another look at cue summation. AV Communication Review, 15 (3), 233-245.
Van der Heijden, M.P. (1987). Tekst en videobeeld. Een onderzoek naar de vormgeving van het videobeeld door teksten en/of live-action videobeelden in informatieve video-voorlichtingsprogramma's [Text images and video images. An experiment into the realization of the video image by texts or by live-action images in informative video programmes]. Enschede: University of Twente, Faculty of Educational Science and Technology, Department of Instrumentation Technology.

Verhagen, P.W. (1993, April). Multi-Media: Basic Approaches for Utilization. Contribution to the COMETT training programme EuroMedia Training: Entwicklung von Fernstudienmodellen im bereich Medizin und Gesundheitswesen für Anwender und Multiplikatoren, Magdeburg, 27-28 April, 1993: Otto von Zernike Universität.

Verhagen, P.W. (1992). Length of segments in interactive video programmes. Unpublished doctoral dissertation. Enschede: Faculty of Educational Science and Technology, University of Twente.

Vincent, W.S., Ash, P., & Greenhill, L.P. (1949). Relationship of length and fact frequency to effectiveness of instructional motion pictures. Technical Report, SDC-269-7-7, Port Washington, NY: Special Devices Center, Office of Naval Research.

Wember, B. (1976). Wie informiert das Fernsehen? (How does television inform?) Munich: List Verlag.

Zuckermann, J.V. (1949a). Commentary variations: Level of verbalization, personal reference, and phase relations in instructional films on perceptual-motor tasks. Technical Report, SDC-269-7-4, Port Washington, NY: Special Devices Center, Office of Naval Research.

Zuckermann, J.V. (1949b). Music in motion pictures: Review of literature with implications for instructional films. Technical Report, SDC-269-7-2, Port Washington, NY: Special Devices Center, Office of Naval Research.