DOCUMENT RESUME

ED 378 955 IR 016 972

AUTHOR Verhagen, Plon W.; Breman, Jeroen
TITLE Instructional Format and Segment Length In Interactive Video Programs.
PUB DATE 8 Feb 95
NOTE 19p.; Paper presented at the National Convention of the Association for Educational Communication and Technology (Anaheim, CA, February 8, 1995).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Comparative Analysis; Data Collection; Higher Education; *Individualized Instruction; Memory; Outcomes of Education; *Performance Factors; Pretests Posttests; *Programmed Instruction; Videotape Recordings
IDENTIFIERS Amount of Invested Mental Effort; *Interactive Videodisks; *Segment Length (Videotape Recordings); Verhagen (Plon W)

ABSTRACT

The purpose of this study was to gather further insight into a previous investigation of the relationship between self-chosen and program-controlled segment length of an interactive videodisk program, and performance on post- and retention tests. The initial study by Verhagen, which questioned what is the optimum length of well-designed audio-visual segments to present factual information via an interactive video program, is reviewed. The additional research reported here was designed to find out whether the conclusions of Verhagen's study should be altered if data collection is carried out in an instruction experiment rather than a memory experiment. It also studied the effect of the amount of invested mental effort (AIME), and tested achievement of instructional objectives concerning the video program as a whole rather than testing specific information elements. Seventy-three university freshmen were given pre-, post-, and retention tests, and were grouped to learn from the following video segmentation formats: (1) Explore--a menu-driven environment giving subject complete control to choose video segments, segment their length, and jump between knowledge questions and remediation video segments at will; (2) Guide--video sequences are pre-segmented, followed immediately by the appropriate knowledge questions and remediation, and presented in linear order; (3) Linear--instructional objectives are presented beforehand, then subjects watch the entire program without stopping; (4) Variable--subject has control over length of each video segment, but not over sequence of video program. Results show that in a memory performance task, shorter video segments were chosen than those chosen for an instructional task. However, no evidence was found that test performance was related to self-chosen segment length. The Explore and Guide conditions required the least mental effort, and thus provided the most relaxed learning situation and usability. Finally, no relationship between AIME and test performance could be determined. Data is presented in nine tables and figures. (Contains 15 references.) (MAS)

Paper presented at the 1995 AECT CONVENTION "Information Technology: Expanding Frontiers", Anaheim, CA, February 8-12, 1995

INSTRUCTIONAL FORMAT AND SEGMENT LENGTH IN INTERACTIVE VIDEO PROGRAMS

Pløn W. Verhagen & Jeroen Breman
University of Twente
Faculty of Educational Science and Technology
Division of Educational Instrumentation
PO Box 217
7500 AE Enschede
The Netherlands.
Email: Verhagen@edte.utwente.nl

Abstract

An experiment is reported in which subjects worked with an interactive videodisk program about cheese making on the basis of the same set of instructional objectives but in different experimental conditions with respect to learner control. The purpose of the study was to investigate the relationship between learner-controlled or program-controlled length of video segments and test performance on post-tests and retention tests. Also, differences between self-chosen and program-controlled segment length regarding the investment of mental effort of subjects were examined. The results show that long program-controlled video segments have some advantage over interactive video when learning factual information. However, well-designed interactive video applications may offer a comfortable work environment that allows for a relaxed use of mental effort. The results of the experiment seem to suggest that video segments with a length of about 3 minutes, containing enough factual information for answering 8 or 9 questions, may appear to be optimal building blocks. Subsequent experiments should reveal to what extent this suggestion is valid.

Introduction

Verhagen (1992a, 1992b) undertook a study to gather insight about the length of segments in interactive video programs on the basis of the following central question:

What is the optimum length of well-designed audio-visual segments to present factual information via an interactive video program to learners who possess certain characteristics?

In his study, five experimental conditions were used in an attempt to answer eleven more elaborate research questions. The study was carried out with a specially produced videodisk program about cheese making. This program contains 252 information elements which form a connected discourse of 36 minutes if the program is played linearly without stopping. An information element is defined as one uninterrupted statement of the off-screen narrator about which one factual question can be put. In two conditions, subjects could choose segment length for themselves by interrupting the video presentation anytime they wanted. In the other three conditions. fixed segment lengths were used to study recall differences with self-chosen segment length. Subjects were 235 first-year students of several technical and social science departments at a single university in the Netherlands.

Verhagen's experiment was essentially a memory experiment in which the program was used to find out what segment lengths are preferred by learners. Each subject worked individually with the program.

The results showed self-chosen segment lengths of which the means per subject varied from 2.19 to 87.50 information elements (which is in time a range from about 16 seconds to 11 minutes). The mean of mean length was 12.70 information elements (which equals about 1.7 minutes) with a standard deviation of 11.77 information elements. Details are reported by Verhagen (1992a, 1992b).

The longer the video segments were, the lower was, on average, the recall of factual information when answering questions directly after watching a segment. The differences between self-chosen segment lengths and fixed segment lengths appeared thereby to be small with the exception of the so-called Linear Condition. The Linear Condition is a fixed-length condition in which the subjects are forced to watch the whole program without stopping. In this condition, the direct-recall performance was substantially lower (54% correct answers) than in any of the other conditions (around 80% correct answers).

After three weeks, delayed recall was measured with a retention test. In this case, subjects of the Linear Condition performed equally well or better than all others while on average questions about long fixed segments were answered better than questions about short fixed segments. The questions about the long fixed segments often also yielded a better performance than the retention scores of subjects who worked with self-chosen long segment lengths. Inspired by work of Salomon (1984) and of Cennamo, Savenye, & Smith (1991), Verhagen (1993) argues that the success of the longer segments may be attributed to larger amounts of invested mental effort (AIME) due to perceived more difficult demand characteristics (PDC's) of the expected longer segments.

Rules about what to regard as 'optimal' when deciding on segment length did not emerge. The findings gave, nevertheless, reasons to formulate a few tentative guidelines for the design of segments. The guidelines that are relevant for the present study are as follows:

> "If factual information is presented to university freshmen and a systematic presentation with off-screen narration is used, a wide variety of segment lengths is feasible. If a safe range is sought in which initial remembering is to reach at least 70%, segment lengths up to 22 information elements are appropriate. Using the condensed presentation rate of the experimental video program used in the study, up to about 3 minutes of presentation time can be used." (Verhagen, 1992b, p.260)

> "For optimal remembering, learners should be strongly urged to invest substantial mental effort. This deepens the initial processing, which enhances long-term storage of the information. Choosing long fixed segments helps, but if the information load goes beyond the ability of the learners they may find the task unpleasant. On the other hand, segments that are too short may be so easy that learners are not compelled to invest effort, and thus hamper adequate storage into memory." (Verhagen, 1992b, p.260-261)

Need for further study

The results in Verhagen's study were obtained through an experimental procedure that follows a pattern of presenting video, asking questions, and giving feedback and remediation. This basic pattern has been used by other researchers who studied interactive video in instruction experiments such as Schaffer and Hannafin (1986), Phillips, Hannafin, and Tripp (1988), and Cennamo, Savenye, and Smith (1991). The procedure can be considered as sufficiently similar to real patterns of instruction to make it possible, in principle, to reach conclusions with respect to segment length, based on the results of the experiments, that can be generalised to rules for the design of segments for instructional interactive video programs. An unsatisfactory aspect of the reported results is, however, that they were obtained in a memory experiment. The performance required did not serve any instructional purpose. The subjects were just expected to recall every factual statement as presented by the program, although literal answers were not necessary (answers to questions that showed understanding of the information that was presented by the pertaining videodisk fragment were rated to be correct). The observed segment lengths may, therefore, not represent the kind of segment lengths that are effective in instructional applications. Looking at other experiments, an obvious difference concerns the information load per segment as expressed in the amount of questions that are presented about the content of each segment. Schaffer and Hannafin (1986), for instance, worked with a video program of 13⅔ minutes that was subdivided into five segments. The mean segment length of their program is 2.75 minutes. The information load per segment was substantially less than in the dense-packed videodisk program on cheese making. Learners had to answer only three questions per segment instead of the 22 questions per segment in the case of segments of 3 minutes in Verhagen's experiment.

Schaffer and Hannafin carried out their experiment with ninth-through-twelfth graders. In the experiment, the amount and type of interactivity was varied in four conditions. The least interactive condition was similar to the Linear Condition of Verhagen's experiment: just presenting the video program without interruption. In the most interactive condition, next to the five video segments, questions, feedback and video branching were components of the interactive structure. Schaffer and Hannafin used a measure that they called "Acquisition Rate". Acquisition Rate was derived based upon the ratio of recall and instructional time. Schaffer and Hannafin found that:
"The fully interactive version yielded the greatest recall but took longer to complete than any of the other presentations. Time was shortest, and the resulting rate of learning was greatest, for the simple linear video presentation" (Schaffer and Hannafin, 1986, p. 89). Schaffer and Hannafin concluded that the efficiency of interactive video for teaching factual content may be questionable. This result seems to show similarities with Verhagen's results, specifically with the relative success of his Linear Condition.

Maccoby and Sheffield (1961), and Margolius and Sheffield (1961) used a film of 18 minutes to carry out a series of studies about segment length and the learning of a mechanical-assembly task. They introduced the idea of "demonstration-assimilation span" (DA-span) which defines a kind of stopping point: "which comes at a point when the learner has assimilated about as much demonstration material as he can hold in his mind well enough to translate into overt practice with little or no error" (Maccoby & Sheffield, 1961, p. 79). Maccoby and Sheffield and also Margolius and Sheffield (1961) used fixed segment lengths. A DA-span segment was defined as a segment which could be practised with a mean group-performance score of about 75 per cent correctly responding immediately after viewing. The length of the segments differed considerably, as the DA-span was determined by how much the student could assimilate in one stretch. One of the four sections of the 18-minute film had a running time of six minutes, whereas another section, a more difficult one, ran only 2.5 minutes, while both were considered to have a similar DA-span for the given target group.
The interpretation of Anderson and Faust (1975, p. 215) of the work by Sheffield and his associates was that the experiments demonstrated "that short presentations followed by short periods of practice are more effective than longer presentations followed by longer periods of practice". They concluded that: "in summary, experimental evidence seems to

indicate that students learn more when required to make active responses immediately after small segments of a lesson" (Anderson & Faust, 1975, p. 216).
The tasks in these experiments were very different from just learning factual information. Students practised the assembly task with concrete parts in their hands and cues from parts already assembled or still on the table. The order of magnitude of the used segment lengths for the assembly task is, nevertheless, rather similar to the segment lengths from the studies by Verhagen and by Schaffer and Hannafin. Although the experimental procedures showed similarities, differences in the tasks, content, and target groups make it impossible to support the validity of Verhagen's first guideline.

This conclusion is reinforced by a study by Weiss, Maccoby, and Sheffield (1961) in which very different segment lengths were in use. Weiss, Maccoby, and Sheffield applied the same research methodology that was used for the assembly task to serial learning of a geometric-construction task. This time a film of 4.5 minutes was used that demonstrated the nine-step construction of a five-sided equilateral polygon, using pencil and paper and a straight-edge and compass. The film consisted of six DA-span segments. The longest DA-span segment in this film was 75 seconds in length, the shortest segment was 12 seconds. They concluded that "the DA-span unit, as the practice unit, uniformly gives superior performance during practice but is not guaranteed to connect together the different parts of the total task. A compromise which gives at least some practice at larger units of the task may be desirable even if it entails more errors in performance at the out-set of practice." (Weiss, Maccoby, & Sheffield, 1961, p. 76)
In the geometry task, all subsequent steps had to be recalled. This requirement resembles more the subject of the other studies mentioned. However, the used segment length were very different. Not much advice can thus be derived from the different experiments about what segment length should be preferred for presenting factual information, possibly with the exception that segments that are too short may hinder the integration of information into one organised whole (Weiss, Maccoby, & Sheffield, 1961). It is, in other words, not clear whether for factual information some frequently alternating presentation-and-practice format should be developed (which complies with the conclusion of Anderson & Faust, 1975) or whether Webb's notion that lengthy films can teach well (Webb, 1975), and may yield a relative efficient learning situation (Schaffer and Hannafin, 1986), should guide designers.

Research questions

The research that is reported here, was designed to gather further insight into that matter and was specifically designed to find out whether the conclusions of Verhagen's study should be altered if data collection is carried out in an instruction experiment instead of in a memory experiment. The second guideline is also involved. Verhagen's contention that the success of the longer segments may be attributed to larger amounts of invested mental effort (AIME), was reached through secondary analysis of his data. In his experiments, no measurement of AIME had been done. In the experiment that is reported here, the study of the effect of AIME is included.
To be able to observe differences with the earlier study, the same videodisk about cheese making is used. The research is undertaken to attempt to answer the next three research questions:

Question 1:
Given an experimental videodisk containing 252 information elements, does self-chosen segment length differ if recall of factual information is not required as an exercise in memory performance with regard to the content of individual information elements but as evidence of reaching explicitly stated instructional objectives concerning the program as a whole?

Question 2:
What differences can be observed between self-chosen and program-controlled segment length regarding the investment of mental effort of subjects?

Question 3:
How do the differences of Question 2 relate to performance on post-tests and retention tests?

Instructional format

The instructional format that had to be developed for the experiments was not very different from the procedure of the earlier experiment. The basic pattern of watching video segments, answering embedded questions and getting feedback on the answers is an appropriate strategy to get active and learning learners, as is demonstrated by research of, for instance, Schaffer and Hannafin (1986), Phillips, Hannafin, and Tripp (1988), and Cennamo, Savenye and Smith (1991). The difference is that the content of the videodisk is analysed to serve a realistic instructional purpose for which clear instructional objectives are formulated. The orientation of the learners takes place on the basis of the objectives and the number of questions have been adjusted to the to-be-learned factual knowledge according to the objectives. The total number of embedded questions in the measuring section of the new version of the program is 97, which is substantially less than the 213 questions in the measuring section of the memory experiment.

The orientation of learners by the use of objectives is a recommended pre-instructional strategy before presenting audio-visual material, as are pre-tests and overviews (Morrison, 1979). The instructional format has been shaped accordingly. A list of objectives was made available to the learners which they were required to study before working with the interactive video program. A pre-test is introduced in the experimental procedure. The control of the program is organised by a menu structure that provides an overview of treated subjects (Figure 1b). This approach is taken despite results by Phillips, Hannafin, and Tripp (1988) who show that cueing by orienting learners with introductory statements before working through video segments is not a very successful way. In their experiment orienting activities of this kind accounted for less than 3% of the variance of post-test scores, while their findings also suggested that cueing by practice is an appropriate strategy. The use of embedded questions may thereby serve as a cueing strategy. In the experiment this kind of cueing is provided also. The orienting activities are not excluded, however, because other research does support Morrison's recommendations. Tovar and Coldevin (1992), for instance, found that the provision of orienting activities in the form of a list of criteria for carrying out a certain procedure, a slide presentation of all the tools used in this procedure, and a flow chart showing the steps of the procedure, significantly facilitated the recall of factual information from an interactive videodisk lesson.

The experimental program

The videodisk program that was used for the study was the same program that was used in Verhagen's experiments. The 252 information elements of this program were divided into two sections. The first 33 information elements form an introduction to familiarise the subjects with the experimental procedure, the remaining 219 information elements form the main program which was used for data collection. In time, the introduction has a length of 4.5 minutes and the main program has a duration of 31.5 minutes. The controlling computer program ensures that the video program can only be started or stopped between two information elements. This causes video segments of the program always to consist of whole numbers of information elements. This mechanism was applied in Verhagen's experiments (Verhagen, 1992a, 1992b) and is maintained in the present study. In the experimental conditions that were developed for the present study, however, the last three information elements of the videodisk program are not used. They provide a not-content-related ending of the linear program that has no function in the menu-driven organisation of the new conditions that are described below.

In the main condition for the present experiment, the measuring section was structured according to the following menu structure:
At the top level of the program, the Main Menu, the subject could choose between the presentation of video sequences or the presentation of questions related to the different sequences (the Main Menu is presented in Figure 1a).
In the Video-Sequence Menu, presented in Figure 1b, the subject could choose one of seven sequences. Once chosen, the video started playing. The subject watched the sequence until he/she decided - for whatever reason - that it was time to command stopping by pushing a mouse button. The program consequently stopped at the end of the information element in which the stopping was requested. Then a menu was presented (Figure 1d) in which the subject could choose to go on with the presentation, to answer questions or to return to the Video-Sequence Menu (Figure 1b). If the subject chose to answer questions he/she was presented a menu (Figure 1c) in which the subject could choose to answer questions about the segment he/she just watched, to answer questions about the entire sequence or to return to the last menu (Figure 1d).
At the end of a sequence the subject was presented a menu (Figure 1f) in which the subject could choose to repeat the sequence, to go to the next sequence, to answer questions or to return to the Video-Sequence Menu. If the option to answer questions was chosen, a menu was presented (Figure 1g) in which the subject had options to answer questions about the last segment he/she watched, answer questions about the whole sequence, go to the Questions Menu (Figure 1c), or to return to the last menu (Figure 1f).

If the subject chose to answer questions about a segment or a sequence, all embedded questions of that segment or sequence where presented successively. All questions were either multiple choice questions or short answer questions. Next feedback was given. In case of a false answer the subject had a second chance to answer the question. This option was built in for the fact that subjects tended to make a lot of typing errors which they could correct at this point. If the answer was still not correct the related video fragment was repeated followed by a third opportunity to answer the question. As soon as feedback was completed, the subject was presented the next menu. After a stop within a sequence the subject could choose to go to the beginning of the next segment (which starts with the information element that follows the one in which he/she stopped) or to return to the Video-Sequence Menu. After a stop at the end of a sequence, the subject could choose to go to the beginning of the next sequence or to return to the Video-Sequence Menu.
If the subject had completed the viewing of a whole video sequence, the related menu item of the Video-Sequence Menu was checked. The same counted for the questions. If all

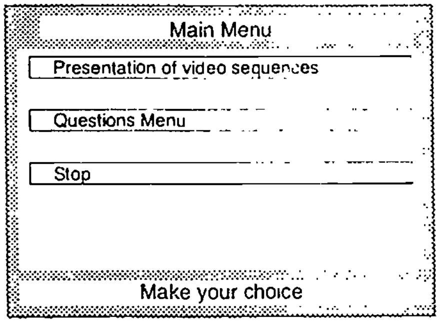


Figure 1a. Main Menu.

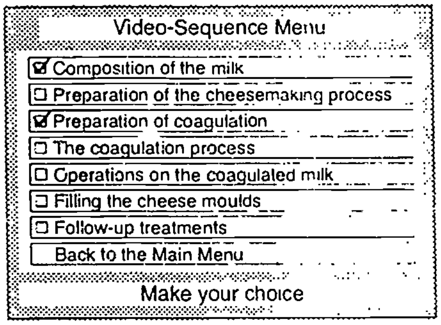


Figure 1b. Video-Sequence Menu.*

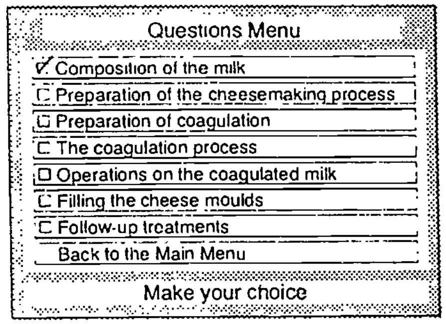


Figure 1c. Questions Menu.**

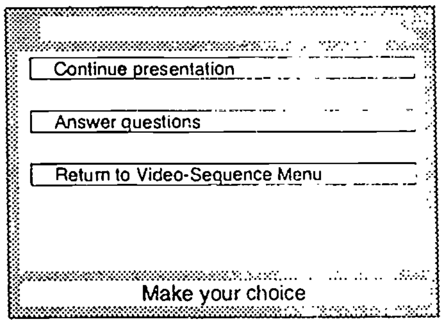


Figure 1d. Menu after a segment not being the last segment of a sequence..

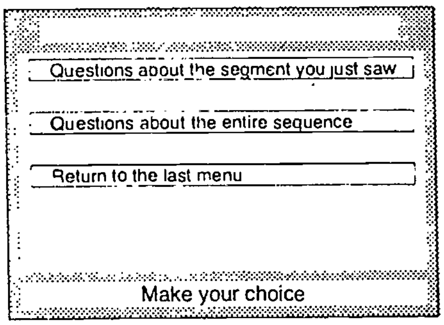


Figure 1e. Questions after a segment not being the last segment of a sequence..

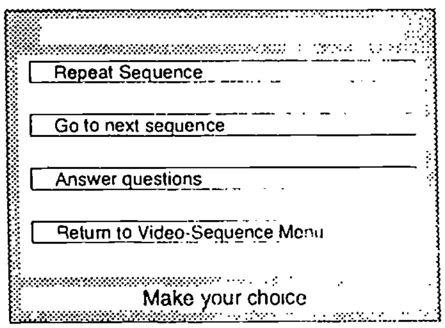


Figure 1f. Menu at the end of a sequence.

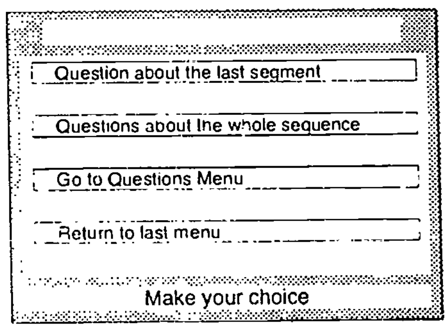


Figure 1g. Questions after the last segment of a sequence.

* The menu in this figure is presented in the situation that the first and third sequence have been watched entirely .

** The questions of the first sequence have been answered.

Figure 1: Menu structure of the experimental program (main condition)

embedded questions of a sequence were answered it was checked in the Questions Menu. The subject had to answer all embedded questions before he/she was allowed to leave the program.

The distribution of information elements, objectives and questions over the different menu items is listed in Table 1.

Table 1. Distribution of information elements, objectives and questions.

Sequences	Number of information elements	Number of objectives	Number of embedded questions
Composition of the milk	25	3	12
Preparation of the cheese making process	27	8	9
Preparation of coagulation	34	7	12
The coagulation process	25	6	20
Operations on the coagulated milk	35	15	17
Filling the cheese moulds	29	4	8
Follow-up treatments	41	7	19
Total	216	50	97

The data presented in Table 1 only account for the main program (Information Elements 34 to 249), which was the measuring section of the program.
The 33 information elements of the introduction -- with an overview of the cheese making process -- counted 25 embedded question. In order to practice the menu structure, the introduction was divided into two sequences.

Subjects

The target group of Verhagen's experiments consisted of university freshmen. This target group is atypical with respect to the educational objectives of the current experiment. Cheese making is mainly taught on different levels of agricultural vocational education. The target group of the current experiments is therefore selected on the middle level of Dutch agricultural vocational education. Two schools participated to yield sufficient subjects. In both schools, first-year students were involved in one of two programs: a three-year program that is more practice oriented (35 subjects) and a four-year program that aims at a higher level of theoretical background than the three-year program does (38 subjects). Analysis of the data revealed that with respect to the current experiment both groups showed no significant differences (Breman, 1995, in Dutch). In this contribution, the research results will therefore be described on the basis of the pooled data of all participants in a certain condition.

Experimental conditions

EXPLORE:
In this condition the subjects were familiarised with the instructional objectives and then offered the menu-driven environment as described before, to explore the videodisk by choosing video sequences from the Video-Sequence Menu (Figure 1b) or by entering the knowledge domain by choosing a group of questions to answer from the Questions

Menu (Figure 1c). The subjects worked through the program, dividing each video sequence if so desired into a series of longer or smaller segments according to the procedure that was explained in the description of the program. The subject could decide to answer, after each segment, the segment-related questions, feedback, and remediation or could choose to do something else (such as interrupting working through a sequence and return to a menu to make a new selection). In this way, a situation of complete learner control existed, with the already mentioned exception that all questions had to be answered before leaving the program.

GUIDE:
In this condition the subjects were familiarised with the instructional objectives as in the Explore Condition and then were offered the different sequences of the program in a linear order. The presented sequences could not be interrupted. The Guide Condition was a fixed-length condition in which presentation of videodisk segments, answering questions, and remediation in case of incorrect answers are incorporated in a tutorial format under program control.

LINEAR:
This condition resembles Verhagen's Linear Condition in that the subjects were forced to watch the whole program without stopping. In this case, however, the subjects were informed about the instructional objectives beforehand just as in the Explore and Guide conditions. Unlike the other conditions, there were no embedded question incorporated in this condition.

VARIABLE:
Given the different target population compared to university freshmen, the main condition of the earlier experiment by Verhagen (1992a, 1992b) was repeated as a fourth condition. In this condition, the subject could operate the program in a way similar to the procedure of the Explore Condition. However, no menu structure was used and no sequences were predetermined. Instead of the 97 objective-oriented questions, the 213 questions of Verhagen's experiment were used. The questions in Verhagen's main program (which was the measuring section) are about the presented factual information that is contained in the information elements. In this condition, the video program was presented in a linear way. The subjects could interrupt the program by clicking a mouse button. The subject then answered all questions about the segment just viewed. The questions were all open questions that required short sentences, single words, or numbers to be typed in as answer. Next feedback was given. Subjects had no chance to correct false answers before remediation. In case of a false answer the related video fragment was repeated followed by a second chance to answer the question and feedback on that answer. After all questions of the segment were reviewed, the subject was allowed to continue the program beginning with the information element that followed the last element of the completed segment. The subject had control over the length of each video segment, but no control over the sequence of the video program.

Research method

The experimental method is a pre-test/post-test/retention-test design.
The data were collected during the September-October period of 1994. Five computer-videodisk-player combinations were available (Olivetti M240 and M290 MS-DOS-computers with MCS Video-overlay boards and Philips videodisk players of the 400 series). The experimental sessions thus had a maximum of five subjects per session. For each subject, written instructions about working with the program were available for

reference, together with a pen and four sheets of paper for making notes, and a study guide that contained the instructional objectives as the main components for the preparation of the subjects to work with the program. The experiment was carried out according to the following procedure:

- The experiment was introduced by a brief oral presentation.
- Next, a paper-and-pencil pre-test of 25 multiple-choice questions was administered.
- This was followed by an introduction of the subject of the experiment by presenting a video segment of 4.5 minutes that gives an overview of the cheese making process. This introductory sequence is composed of the first 33 information elements of the videodisk program.
- In the learner-control conditions the experimental procedure was then practised by working through the introductory sequence a second time in the way that was required for the measuring section (the 219 information elements of the main program). In the other conditions this step was skipped.
- In the next step, the subjects were invited to study the instructional objectives that were presented in a study quide.
- A questionnaire was administered with respect to the expectancies related to working with the experimental program. The subjects answered three five-point-scale questions to get an indication of their attitude towards the program (I expect to like/dislike working with the program), the Perceived Demand Characteristics (PDC) of the program (What I have to learn seems difficult/easy to me), and the Perceived Self Efficacy (PSE) with respect to the task (To be able to complete the post-test satisfactory, I will have to invest much/little effort during working with the program).
- After all these preparatory activities, the experimental task was carried out. The researcher remained available for answering questions and solving mainly minor technical problems that did occasionally occur.
- After completion of the task, the subjects were allowed to go one time more through their notes and the list of instructional objectives.
- The next step was administering a questionnaire of seven or eight five-point-scale questions (dependent on the experimental condition). The first three questions were similar to the questions of the first questionnaire. The question "I liked/disliked working with the program" was used to find out whether the initial attitude did change. The question "What I had to learn was difficult/easy" was again intended to get an indication of the PDC of the program. The question "I had to invest much/little effort during working with the program" was used to get an indication of the Amount of Invested Mental Effort (AIME) to complete the experimental task. The other questions mainly concerned the audio-visual presentation of the program. The answers revealed no problems with the quality of the program as a professional audio-visual product. These questions will therefore not be discussed further. More information can be found in Breman (1995, in Dutch). One question asked to estimate the expected success in making the post-test (Now that I completed the task with the program, I expect that I will perform well/poorly on the post-test).
- After the questionnaire, a paper-and-pencil post-test of 25 multiple-choice questions was administered.
- After the post-test, a third and last questionnaire was administered, this time consisting of four five-point-scale questions. One question concerned the test performance (Now that I completed the post-test, I think that I did well/poorly). This question relates to the question about the expected success in making the post-test. The results show that the subjects made a poor estimate of their ability to perform well on the post-test, while their judgement of their performance on the post-test after completing it, was much better. For the pupose of the present study, this issue needs no further attention. The other questions asked for preferences with respect to learning from books, video, or from the teacher. The results of these questions are, again, not included in the present contribution. Detailed results are available from the authors.

The post-test concluded the first session.
- After three weeks a paper-and-pencil retention test was administered that again had 25 multiple-choice questions.

Pre-test, post-test and retention test were composed from a pool of 75 questions that consisted of 25 groups of three questions. The three questions in a group were variants of one and the same question. To prevent test bias, the allocation of each of the three questions in a group to either the pre-test, the post-test, or the retention test was done randomly for each individual subject. Each subject thus received own versions of the three tests.

Results

The results of the study are as follows.

Self-chosen segment length

Table 2 and Figure 2 show the differences between the Variable Condition and the Explore Condition with respect to self-chosen segment length.

The mean of mean self-chosen segment length in the Explore Condition appears to be more than twice the length of the mean self-chosen segment length in the Variable Condition. This difference is significant ($N=33$, $F=23.68$, $p<.0001$, tested with a one-way analysis of variance).
A possible explanation could be that the number of embedded questions was very different in these two conditions. To be able to reflect on that difference, segment length is also expressed in terms of the number of questions that the subjects had to answer for each segment. Mean and standard deviation (SD) of this measure are presented in the two most-right columns of Table 2. In this case, no significant differences are found ($N=33$, $F=0.22$, $p=.64$, tested with one-way analysis of variance). The interpretation of this fact is, however, difficult because of the unequal distribution of embedded questions and learning objectives over the 7 video sequences of the Explore Condition (Table 1).

Table 2: Mean of mean self-chosen segment length by condition.

		Mean segment length expressed in terms of			
		number of information elements		number of questions	
Condition	N	M	SD	M	SD
Explore	20	21.55	7.75	9.50	3.37
Variable	13	8.85	6.59	8.69	6.46

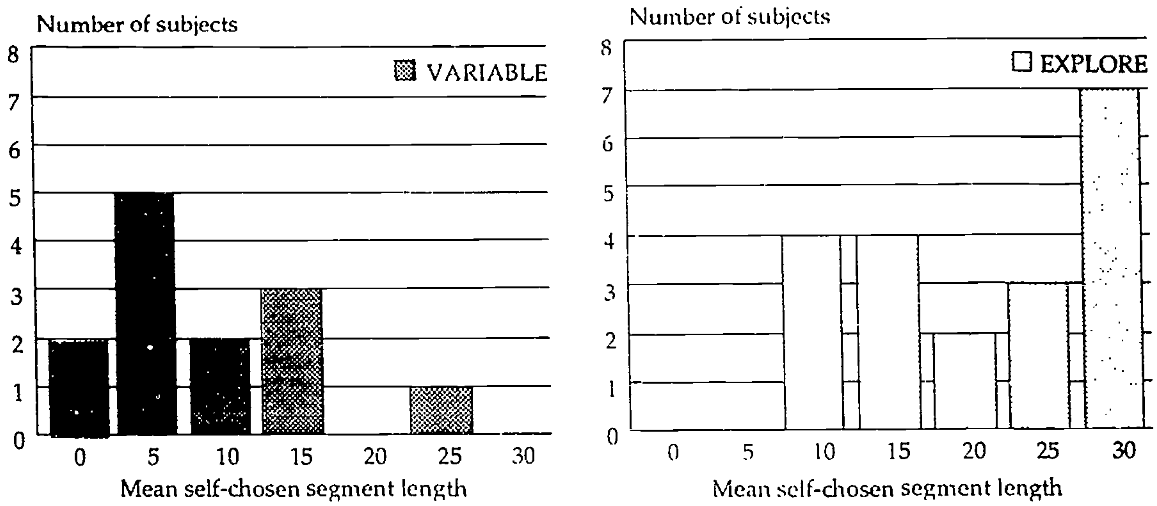


Figure 2: Distribution of self-chosen segment length

Test performance and self-chosen segment length

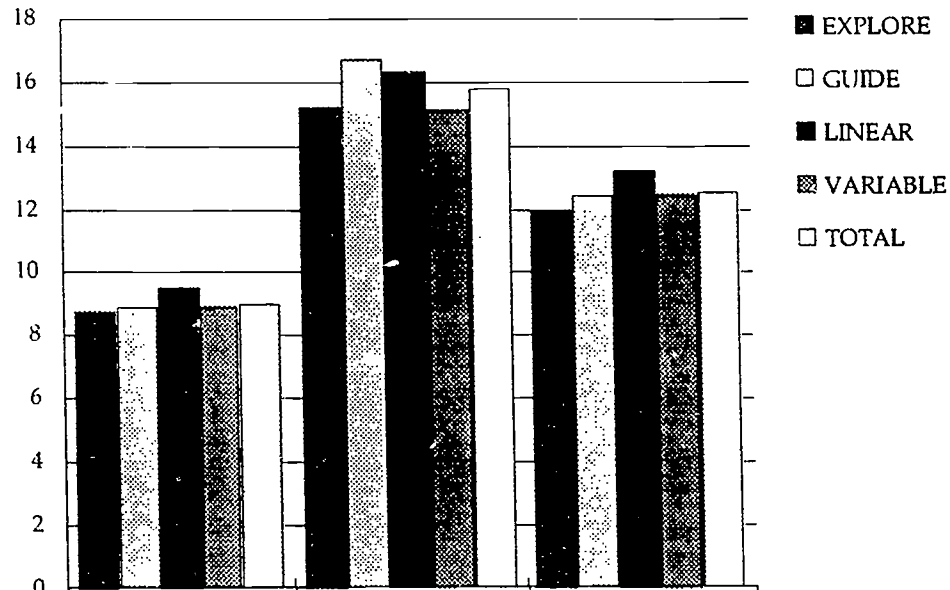
The relationship between mean self-chosen segment length and test performance is analysed by correlating mean self-chosen segment length with the differential scores of the different tests. Table 3 shows the correlations of the Variable Condition as well as of the Explore Condition with the differences of post-test and pre-test scores and of retention test and pre-test scores. For this analysis the non-parametric correlation coefficient 'Kendall's tau' has been used. Table 3 shows that there is no reason to infer that mean self-chosen segment length affected test performance.

Table 3: Correlation between mean self-chosen segment length and test performance

		Post-test minus pre-test scores		Retention test minus pre-test scores	
	N	τ	P	τ	P
Condition:					
Explore	20	0.07	0.74	-0.04	0.80
Variable	13	-0.03	0.89	-0.20	0.34

Test performance and condition

Figure 3 shows mean test scores on pre-test, post-test and retention test for each of the experimental conditions. Test scores are absolute scores. The maximum test score was 25 in all cases.
No significant differences are observed with respect to pre-test scores (N=73, F=.22, p=.88). For the post-test and retention-test scores, the data of Figure 3 are not suitable to analyse differences between conditions. The reason is that pre-test scores varied considerably, ranging from 3 to 16.



Condition	N	Pretest M	Pretest SD	Posttest M	Posttest SD	Retention test M	Retention test SD
EXPLORE	20	8.75	3.14	15.25	4.04	12.00	3.10
GUIDE	18	8.94	2.48	16.72	4.03	12.50	2.81
LINEAR	18	9.50	3.03	16.33	2.00	13.24	2.02
VARIABLE	16	8.94	3.07	15.15	3.63	12.46	2.88
TOTAL	72	9.03	2.90	15.90	3.52	12.53	2.72

Figure 3: Mean test scores per condition

The assumption that the subjects had no pre-knowledge about cheese making appears thus not to be completely true. To estimate knowledge gain, differential scores are used. Figure 4 shows the mean scores of post-test minus pre-test and of retention test minus pre-test for each of the conditions.

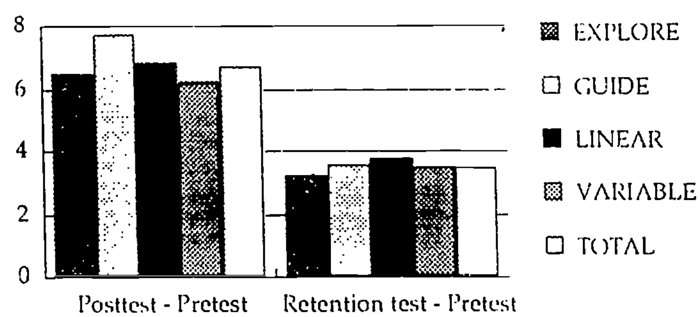


Figure 4: Mean differential scores per condition

Figure 4 shows that subjects who worked in the Guide Condition gained most, on average, between pre-test and post-test. Subjects in the Linear Condition did best on the retention test. Two-way analysis of variance, however, revealed no significant differences.

AIME

The score on the question "I had to invest much/little effort during working with the program", which was part of the questionnaire that was administered directly after the completion of working with the experimental program, was used as an estimate of the Amount of Invested Mental Effort (AIME). The score on the question "To be able to complete the post-test satisfactory. I will have to invest much/little effort during working with the program", which was part of the questionnaire that was administered before working with the experimental program. was used as an estimate of Perceived Self Efficacy (PSE). The Perceived Demand Characteristics (PDC) of the program were rated twice: by the score on the question "What I have to learn seems difficult/easy to me" before working with the program and by the score on the question "What I had to learn was difficult/easy" after working with the program. Table 4 shows the correlations (Pearson's R) between AIME and the other measures. The significant results for the relationship between AIME and PSE suggest that the subjects were honest in their answers and indeed seriously invested effort when they expected that the difficulty of the task needed it. This is reinforced by the fact that more significant positive correlations emerged between AIME and PDC after working with the program than before working with the program. The experience of working with the program seems to have helped the learners to adjust their AIME to the task.

Table 4: Correlations of AIME with PDC and PSE.

AIME versus:	PDC before		PSE		PDC after	
	R	P	R	P	R	P
ALL SUBJECTS	0.39	<0.001	0.68	<0.001	0.48	<0.001
EXPLORE	0.55	0.01	0.77	<0.001	0.44	0.05
GUIDE	0.18	0.47	0.61	<0.01	0.68	<0.01
LINEAR	0.34	0.16	0.79	<0.001	0.59	0.01
VARIABLE	0.21	0.45	0.52	0.04	0.00	1.00

Table 5 shows that the subjects adjusted their PDC due to working with the program. In three conditions, a significant positive correlation seems to imply that the initial PDC was in most cases a fair estimate. In the Variable Condition this was clearly not the case. The reason could be that practicing the experimental procedure with the introduction was eventually not perceived as representative for the task in the measuring section. The data showed that in this condition. PDC was rated to be slightly more difficult before working with the program than after working with the program.

Table 5: Correlations of PDC before working with the program with PDC after working with the program.

PDC after versus:	PDC before	
	R	P
ALL SUBJECTS	0.46	<0.001
EXPLORE	0.73	<0.001
GUIDE	0.56	0.02
LINEAR	0.46	0.06
VARIABLE	0.04	0.89

The effect of AIME on test performance was analysed using a Kruskal-Wallis one-way analysis of variance (Table 6). No significant relationships were discovered. In the Explore Condition a light tendency towards some influence of AIME is visible.

Table 6: Differences in performance by AIME scores for the different conditions.

	AIME scores versus:			
	post-test minus pre-test scores		retention test minus pre-test scores	
	Chi^2	P	Chi^2	P
ALL SUBJECTS	3.13	0.37	1.37	0.71
EXPLORE	5.35	0.15	5.90	0.12
GUIDE	0.36	0.95	2.13	0.55
LINEAR	2.53	0.47	3.13	0.37
VARIABLE	0.36	0.83	1.05	0.59

The difference of AIME-Scores between condition showed some significance (Table 7). Analysis of the data, using the Mann-Whitney U-Wilcoxon Rank Sum W test corrected for ties, showed that for working with the Explore Condition significant less mental effort was invested than in the Linear Condition (P=.05) or in the Variable Condition (P=.01). Other relationships were not significant, although the Guide Condition shows a tendency to have required less effort than the Variable Condition (P=.09).

Table 7: Difference in AIME-Scores between conditions.

	Conditions	
	Chi^2	P
AIME	8.07	0.05

Discussion

In a first instance, the mean of mean self-chosen segment length in the Variable Condition (mean of mean length: 8.85) seems to differ from the results from Verhagen's earlier experiment. However, the mean of mean segment length in his study (12.70 Information

Elements) was in itself the mean of the results from three different groups. They were university students from engineering department (mean of mean length: 17.66 Information Elements), student from a faculty of public administration (mean of mean length: 14.07 Information Elements), and students form an faculty of educational science and technology (mean of mean length: 7.42 Information Elements). The results of the present study compare to the results of the last group and can thus be regarded to fall within the normal range of the experimental method of the earlier study.
The difference between the Variable Condition and the Explore Condition seem therefore to be attributable to task differences. Although the subjects in the Variable Condition were oriented on the experimental task by the pre-test and the list of learning objectives just as the subjects in the other conditions were, it seems likely that the large number of very specific questions in thi condition did shape their behaviour. They adopted, in other words, a strategy for working through the program that compares to the university students in the earlier experiment by Verhagen (1992a). The result that AIME and PDC showed hardly any or no relationship in the Variable Condition (Table 4 and Table 5) may be interpreted to support this view. In all other conditions, the task requirements were consistent with the orienting activities. In these conditions, embedded questions were designed to relate to the explicitly stated learning objectives. Working with the program thus helped to clarify PDC. The own set of questions of the Variable Condition and the relatively rigid experimental procedure in this condition may have caused interference in that respect.
The answer to the first research questions seems to be positive. The task structure of the Variabie Condition, which was in essence an exercise in memory performance, lead to segment lengths that were, on average, four-tenths of the length of the self-chosen segments in the Explore Condition, which was an instructional environment under learner control.

No evidence was found that test performance was related to self-chosen segment length (Table 3). Differences in test performance between conditions were small. Subjects who worked in the Guide Condition did best on the post-test. Subjects who worked in the Linear Condition did best on the retention test. These results are similar to the results of Verhagen (1992b). Mean length of the fixed segments in the Guide Condition was 30.86 Information Elements (in time: about 4,5 minutes), while the mean of mean self-chosen segment length in the Explore Condition was 21.55 Information Elements (in time: about 3 minutes). Also Verhagen found that forced watching of relative long segments supports learning. Whether this implies more invested mental effort, however, is not clear. The self-reported AIME by the subjects does not show any relationship with test performance (Table 6). With respect to the Linear Condition: In Verhagen's case this condition also outperformed all conditions on the retention test. He suggested that the fact that subjects who watch the linear version are provided with an overview of the whole program, may support later recall. The data from Breman (1995) show that in the Linear Condition the subjects completed the experimental task in about half of the time that was needed for the Explore Condition or the Guide Condition. The Variable Condition took even more time. Breman further found that after working with the program, the subjects in the Linear Condition liked the program better than the subjects in any of the other conditions. These facts, combined with the positive results of the Linear Condition regarding retention-test performance, suggest that when video is considered for instruction with the acquisition of factual knowledge as its main goal, linear video should be preferred instead of interactive video. This conclusion is further supported by the conclusion of Schaffer and Hannafin (1986) who found that linear video had the highest Acquisition Rate.

The second research questions asks what differences can be observed between self-chosen and program-controlled segment length regarding the investment of mental effort of subjects. The answer is found in Table 7. According to the self-reports by the subjects, for working with the Explore Condition significant less mental effort was invested than in the Linear Condition or in the Variable Condition. The Guide Condition shows a tendency to have required less effort than the Variable Condition. The Explore Condition and the Guide Condition are new conditions that have been developed to provide a well-designed learner-controlled or program controlled system, respectively. Both conditions appear to offer the most relaxed learning situation and -- if comfort may be taken as a criterion -- the best usability. This usabilty aspect does not support the preference for linear video that was concluded before.

That the Guide Condition requires more AIME than the Explore Condition is probably due to the fact that the mean length of the fixed segments in the Guide Condition was about 1,5 times the mean of mean self-chosen segment length in the Explore Condition. It would be interesting to know how data would change if the mean length of the fixed segments in the Guide Condition would be adjusted to be similar to the mean length of the Explore Condition. Would AIME be reduced? Would then the better performance on the post-test also be reduced? A clear recommendation which of these conditions should be preferred is, in any case, not possible on the basis of the present data.

The third research question concerns the relationship between AIME and test performance. No such relationship could be demonstrated (Table 6).

Conclusion

Different approaches for segmenting video and organising interaction appear to yield different results with respect to test performance, instruction time, AIME, and liking of a learning task that concerns the acquisition of factual knowledge. Linear video shows some advantages for learning of factual knowledge. A clear preference for linear video over interactive video is, however, not indicated in all circumstances. On the basis of the results of the present experiment, the amount of mental effort that is required for relatively long linear programs seems to be more than in interactive programs in which the video is segmented in a proper way. In situations where next to learning factual knowledge also more complex learning outcomes are desired, the advantages of linear video may remain valid for the factual components of the instruction. A well-designed interactive video application may offer a comfortable work environment that allows for a relaxed use of mental effort. The results of the experiment seem to suggest that video segments with a length of about 3 minutes, containing enough factual information for answering 8 or 9 questions, may appear to be optimal building blocks. Subsequent experiments should, however, reveal to what extent this suggestion is valid.

References

Anderson, R.C., & Faust, G.W. (1975). *Educational psychology, the science of instruction and learning*. New York: Dodd, Mead & Company.

Breman, J. (1995). Het effect van instructieformat op segmentlengte bij instructie met interactieve video [The effect of instructional format on segment length in interactive video lessons]. Enschede: University of Twente, Faculty of Educational Science and Technology.

Cennamo, K.S., Savenye, W.C., & Smith, P.L. (1991). Mental effort and video-based learning: The relationship of preconceptions and the effects of interactive and covert practice. *Educational Theory, Research, & Development*, *39* (1), 5-16.

Maccoby, N., & Sheffield, F.D. (1961). Combining practice with demonstration in teaching complex sequences: Summary and interpretation. In A.A. Lumsdaine (Ed.), *Student response in programmed instruction* (pp. 77-85). Publication 943. Washington, DC: National Academy of Sciences - National Research Council.

Margolius, G.J., & Sheffield, F.D. (1961). Optimum methods of combining practice with filmed demonstration in teaching complex response sequences: Serial learning of a mechanical-assembly task. In A.A. Lumsdaine (Ed.), *Student response in programmed instruction* (pp. 33-53). Publication 943. Washington, DC: National Academy of Sciences - National Research Council.

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.

Phillips, T., Hannafin, M., & Tripp, S. (1988). The effects of practice and orienting activities on learning from interactive video. *Educational Communications and Technology Journal*, *36*, 93-102.

Salomon, G. (1984) Television is easy and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, *76* (4), 647-658.

Schaffer, L., & Hannafin, M.J. (1986). The effects of progressively enriched interaction on learning from interactive video. *Educational Communications and Technology Journal*, *34*, 89-96.

Tovar, M., & Coldevin, G. (1992). Effects of orienting activities an instructional control of learning facts and procedures from interactive video. *Journal of Educational Computing Research*, *8* (4), 507-519.

Verhagen, P.W. (1992a). *Preferred length of segments in interactive video programmes*. In M.R. Simonson & K. Jurasek (Eds.), 14th annual proceedings of selected research and developments presentations at the 1992 National Convention of the Association for Educational Communications and Technology (AECT), Washington DC., February 5-9, 1992. Ames (Iowa): Iowa State University, College of Education, Instructional Resources Center.

Verhagen, P.W. (1992b). *Length of segments in interactive video programs*. Doctoral dissertation. Enschede: Faculty of Educational Science and Technology, University of Twente.

Verhagen, P.W. (1993, August). Amount of invested mental effort and length of video segments in interactive video programs. Paper presented at the 5th EARLI Conference, Aix-en-Provence.

Webb, J.E. (1975). Learning from a film. *Journal of College Science Teaching*, *5* (2), 138-141.

Weiss, W., Maccoby, N., & Sheffield, F.D. (1961). Combining practice with demonstration in teaching complex sequences: Serial learning of a geometric-construction task. In A.A. Lumsdaine (Ed.), *Student response in programmed instruction* (pp. 55-76). Publication 943. Washington, DC: National Academy of Sciences - National Research Council.