

# Video segmentation as an example for elaborating design patterns through empirical studies

Niels Seidel

niels.seidel@fernuni-hagen.de

Chair for Cooperative Systems at FernUniversität in Hagen  
Hagen, Germany

## ABSTRACT

While pattern mining is characterized by empirical examinations the applied solutions propagated in the patterns are only rarely the subject or the result of empirical tests. However, confidence that a pattern will serve its purpose can only be obtained through testing and using them in practice.

This article uses the pattern VIDEO SEGMENTATION as a use case to present an iterative approach on how to integrate iterative empirical tests into the elaboration of design patterns. This approach aims to explore the design space and specifies the context of a pattern. With the approach presented here, we want to encourage researchers from different disciplines to communicate their empirical research results in the form of design patterns.

Using learning analytics methods, two studies were conducted for this purpose. In the first study 10 popular educational YouTube channels were analysed regarding the length and sequence of 4.136 videos. By using a cluster analysis three groups of videos of different length could be identified: videos under 8 minutes playing time, videos between 8 and 20 minutes duration and longer videos.

For the second study 22 participants were split up into two groups using a video player supporting segmented videos in two different ways, and a control group watching a non-segmented video. Segmented videos with sequences of up to 10 minutes resulted in higher learning effects than the non-segmented version of the same video.

## CCS CONCEPTS

• **Human-centered computing** → *Social media*.

## KEYWORDS

design pattern theory, learning analytics, video segmentation, pattern mining

## ACM Reference Format:

Niels Seidel. 2020. Video segmentation as an example for elaborating design patterns through empirical studies. In *European Conference on Pattern Languages of Programs 2020 (EuroPLoP '20)*, July 1–4, 2020, Virtual Event, Germany. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3424771.3424778>

## 1 INTRODUCTION

The context of a design pattern describes the application domain and the conditions under which the solution described in the pattern works. If there are a lot of alternative applications domains or conditions mentioned in the context statement, the pattern claims a lot. It therefore seems logical, that a “pattern should only claim as much as has been tested” [17, p. 139]. In that sense patterns are considered as testable theories [17, p. 129] or hypotheses. A passed test provides confidence in the hypothesis, but a finally failed test requires a rejection of the hypothesis.

In the practice of pattern mining and writing testing plays a minor role. The *Rule Of Three* suggests to find known uses to prove the proposed solution. The way of “proving” is meant rather as observing, tracing back, or describing examples, instead of scientific rigor in testing the solution under different conditions. It becomes even more complicated when it is recognized that design patterns do not occur in isolation, but appear as a network of patterns. As a consequence, patterns cannot be tested separately without considering each force. However, the existence of multiple interdependent variables is not a reason not to test. In the social sciences and in psychology, for example, the multivariate analysis of variance (MANOVA) is used for this purpose.

A further difficulty resulting from the *Rule Of Three* is the compulsion to derive the essence of a solution from existing (hardly validated) examples instead of being able to present well-founded and empirically tested developments as a patterns. This compulsion is one reason why patterns seem hardly applicable to practitioners and scientists of disciplines like psychology, learning sciences, but also subfields of computer science [44].

From the perspective of scientific theory the pattern elaboration process incorporates parallels to design-science research [16] as well as design based research [12, 20]. Thinking the way of design-science patterns can on the one hand be considered as viable and relevant artefacts that are iteratively processed in the “cycles of design, testing, analysis, and redesign” [12] (cf. [38]). On the other hand patterns can be

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*EuroPLoP '20, July 1–4, 2020, Virtual Event, Germany*

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-7769-0/20/07...\$15.00

<https://doi.org/10.1145/3424771.3424778>

viewed as the resulting principles of an applied design based research process.

This article uses the example of the pattern "Video Segmentation" to present an iterative approach on how to combine the elaboration of design patterns with the conduction of empirical studies. It shows ways to explore the possible application domains and the design space of a pattern.

This paper contributes to the pattern theory literature by presenting a process model for the refinement of design patterns based on empirical studies. We are particularly interested in how, in addition to the known issues, comprehensible empirical studies enable a more precise delimitation of the context and the solution. Another contribution in the area of Learning Analytics is to analyse the structure and extent of videos in relevant YouTube channels. With regard to the design of video learning environments, we make a contribution in the field of video-based learning. We conducted a user study (N=24) about two treatments for presenting two types of segmented videos in a interactive video player. For practitioners in the field of technology-enhanced learning the presented design pattern VIDEO SEGMENTATION may support to the design of video learning environments and the production of video-based learning resources.

The remainder of the paper is structured as follows. In the next section we discuss related work. After that the process of empirical elaboration of design patterns is presented in section 3. In section 4 we take a look at existing theories and the state of empirical research on the length and segmentation of learning videos. Section 5 is dedicated to two empirical studies. In the first study the current practice in the production and application of educational videos is examined. The second study is a user study, in which design variations in relation to the realization of the video segmentation are examined comparatively. The pattern VIDEO SEGMENTATION as the preliminary result of the applied pattern elaboration process presented in section 6. The paper ends with a conclusion and an outlook in section 7.

## 2 RELATED WORKS

Seidel [40] focused on video-based learning environments incorporating the design of video players and platforms that host collections of videos. Although patterns like SEQUENTIAL MEDIA describe the technical foundation the matter of video segmentation as a possible step in the video production process was not covered in that pattern language. Koumi [19], Bates [3], and Schön and Ebner [34] worked out more or less uniformly structured manuals and guidelines for producing educational videos. Further production related patterns dedicated to screencast was presented by Chen and Rabb [8]. In these works segmentation is not seen as an option in post-production or as a design feature of a video player.

In another production-related pattern Mor and Warburton [28] emphasized the need for limiting the video length in the pattern SIX MINUTE VIDEO. In relation to the objectives of this paper, this pattern refers to a very limited design space (video length) and covers an equally limited

application domain (MOOCs). In the sense of the previously presented iterative process for the empirical elaboration of design patterns, the pattern SIX MINUTE VIDEO can serve as a starting point for empirical-based improvement.

In educational psychology, a number of empirical findings on the segmentation of dynamic visualizations have been shared with the scientific community. In section 4 we will discuss this in more detail. However, many of these studies consider isolated variables under controlled conditions in the laboratory. However, there are only a few field studies. In contrast, in the field of Learning Analytics almost exclusively empirical data are collected and analyzed in the field. Guo et al. [14] and Li et al. [21], for instance, have developed pattern-like design recommendations based on studies of user behaviour in video-based Massive Open Online Courses (MOOCs). Empirical well-founded principles for multimedia learning were proposed by Mayer [24] and Clark et al. [9]. Although these recommendations, principles, and guidelines have a practical use, the reference to the design of adequate software applications is missing. Compared to the format of design patterns, they lack structure.

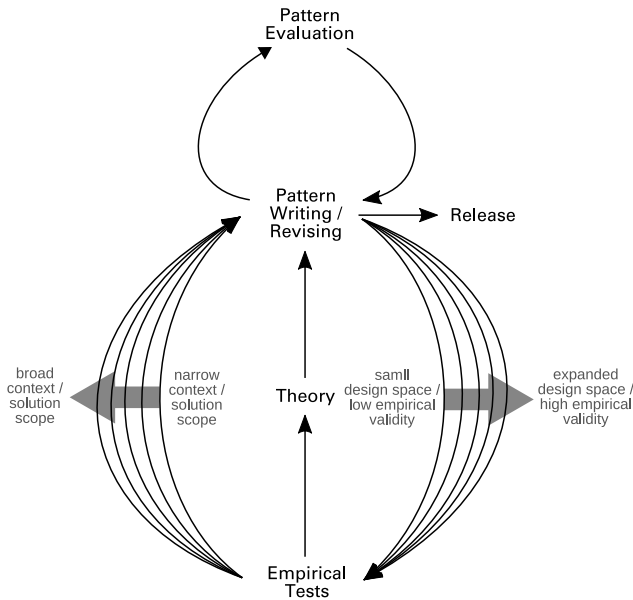
Empirical studies also play an important role in pattern mining. For example, artefacts (e.g. literature) are evaluated in which experts reflect their experiences, comparisons with similar systems are made (cf. [37]), ethnographic observations are conducted (cf. [2, 22, 31] and [35, p. 85-91]) in addition, self-observations in the form of retrospectives are carried out (see [2]) or experts are asked about their own experiences and self-identified patterns (see [1, 5]). A more detailed overview of the different methods can be found at Kohls [17, p. 135f] as well as at Seidel [38]. It is remarkable, however, that the application of the solution propagated in the pattern is no longer the subject of empirical studies. If design patterns represent hypotheses according to Kohls [17, p. 129], we can observe a lack of attempts to test these hypotheses in practice.

## 3 PROCESS OF EMPIRICAL ELABORATION OF DESIGN PATTERNS

The *Irsee onion dom*<sup>1</sup> model shown in Fig. 1 represents the process cycles on how empirical studies fertilize the pattern writing and revision.

The model consists of a lower and an upper loop. The upper loop symbolizes the evaluation and revision of the written manifestation of a design pattern. Typically, this iterative process will take place in a circle of pattern authors. Also feedback processes like the *Shepherding* [15] and the *Writers' Workshops* [10], as they are common on pattern conferences, lead to multiple evaluations and revisions of a pattern text. This applies analogously to Pattern Languages Seidel [38]. However, the subject of this process loop are formal criteria of the pattern format, the comprehensibility of the language and the plausibility, which is verified through the experience

<sup>1</sup>The name is an allusion to the typical shape of Bavarian church towers, as they can be found at the Irsee Monastery, where the EuroPLoP takes place.



**Figure 1: The *Irsee onion dom* representing the process of empirical elaboration of design patterns**

of the experts involved. The actual subject of the pattern solution is not subject to any test that could meet strict scientific criteria. For this reason, we argue that a second loop is needed in which empirical data is collected, evaluated and used to revise the pattern.

The lower loop in the process model of Fig. 1 is based on the research and development cycles of Euler [13]. The design of a prototype or a procedure is refined and empirically investigated according to a particular research design. The findings of the investigation are used to revise the pattern and simultaneously substantiate statements of the pattern. The core idea is to stepwise extend the design space or the empirical validity of a study in each cycle in order to expand the context of the pattern or to specify the solution if the tests were successful. If the tests fail and the proposed hypotheses have to be rejected the context of a pattern needs to be narrowed or the scope of a solution needs to be adjusted. In an extreme case a design pattern has to be withdrawn. This gradual expansion makes it impossible in most cases to use existing empirical research. Existing works often have a slightly different scope and do not expand the design space in a systematic manner.

This described process can be started at three different points within the two loops:

- Entry point 1: Based on empirical research, generalizable principles and principles are derived and formulated in the form of a design pattern. Further empirical investigations are carried out to determine the extent to which further generalization is possible. Apart from this, the written representation of the pattern is iteratively revised in the manner described above.

- Entry point 2: Starting from a given pattern, validity in the described design space is empirically examined. If the evaluation is successful, the design space can be extended or the scope of the solution can be expanded for the next testing round. If this test is also successful, the results will be formulated accordingly in the context and the solution statement of the pattern. Starting from the new version of the patterns
- Entry point 3: A special case of entry 2 exists, if proven theories form the foundation of a pattern. The empirical investigation of a design artifact then serves not only to verify the statements in the pattern, but also to test hypotheses that may support or falsify a theory (and thus the pattern based on it).

As a consequence of this approach, design pattern become fluid instead of manifested and static. If a pattern should be published – as for example in this article – it is only a representation of the current available knowledge about the subject, which is based on empirical investigations and was brought into a comprehensible written structure considering the feedback from colleagues and other experts. In the sense of Kohls [17] a pattern is to be regarded as a theory that is subject to further investigations and tests [18].

By applying this procedure we want to give an impulse to revise and empirically test existing patterns. This will increase the scientific importance of patterns as a profound knowledge representation and thus their relevance for the dialogue between science and practice.

In the further course of this paper we are using the pattern SIX MINUTE VIDEO [28] as an example to demonstrate the application of the process described above. Therefore, we start at the 2nd entry point and continue with two empirical studies related to the length and segmentation of learning videos. Before that, we will look at the theoretical and basic principles and the empirical studies that are available on this subject.

## 4 THEORETICAL FOUNDATIONS AND EMPIRICAL RESULTS FOR VIDEO SEGMENTATION

Short videos are preferred by user, but what is short? The video platform YouTube itself classifies videos as “short” if they are under 4 minutes long. “Long” videos, on the other hand, are all over 20 minutes. Short videos are also preferred by students. The short length of the videos is intended to prevent the viewer from becoming too passive, and to prevent thoughts from wandering [7, 32]. Guo et al. [14] analysed the engagement time (video watching session length) in multiple MOOCs by comparing videos of different type (e.g. lecture, tutorial, other), production style (e.g. slide presentations, classroom, Khan-style), and video length. At least for MOOC-like setting Guo et al. [14] recommended that individual videos should be kept shorter than 6 minutes, except tutorials where the duration does not matter. However, this results can not be generalized for all types of video, production styles and learning scenarios. For an in-depth analysis a comprehensive

classification of learning videos such as presented by Seidel [39] should be taken into account.

In any case, *shorting* is a design decision for which many implementation possibilities exist. During video production, content can be streamlined by eliminating side issues, repetition of pre-knowledge (cf. *pre-training principle* [24]) or redundant informations (cf. *redundancy principle* [26]). Increasing the speaking rate will save a few seconds, but can lead to more attention according to Guo et al. [14]. Content can be also outsourced to other videos by referencing them. In post-production, pauses in speech can be removed [14], the playback speed can be increased by a factor of two and visual content can be reduced to relevant still images. If a video is compressed in the described way, it may have to be watched several times. A compressed learning video is therefore not necessarily self-contained, since it is not possible to thoroughly examine the learning content without dealing with additional learning materials. Last but not least, it should not be forgotten that passages with lower information density give viewers a break. If a video is reduced to the relevant and densely packed passages, the viewer has hardly any time left to breathe.

As an alternative to shorting videos, the *segmenting principle* postulates the division of longer passages into smaller, separate sections. After the individual video, the learner usually has to become active himself and click on the next video [7]. This can be done by segmentation during post-production or by better planning of the videos with regard to their length during pre-production. Segmentation is said to have two main advantages that have a positive impact on learning facilitation: a reduction in cognitive load and better structuring of the learning material [41].

Mayer and Chandler investigated the effect of segmentation on learning by letting learners watch an animation of the creation of lightning. One group was shown a version in which the entire learning content was presented as a unit and a second group was shown a version with 16 individual segments [25]. In a subsequent transfer test, the learners who had worked with the segmented version and who had to actively click to the next section in each case performed significantly better. In addition, a further experiment tested in which order a holistic video and a segmented video should be used. The “Whole-Part” approach describes the showing of the whole video and then the segmented individual sections. The “Part-Whole” approach describes the display of the individual segmented sections and then the entire video. Congruent to the theory of cognitive load, the learners of the “Part-Whole” group scored better [25].

These advantages of segmentation shown by Mayer and Chandler were confirmed by Moreno [29] a few years later. In Moreno’s study, prospective teachers were shown videos about principles or techniques of teaching. A control group was shown a continuous 20-minute video with the teaching techniques. A test group was presented with the same content in seven segmented individual videos and each video presented only one technique. Again, the learners who saw

the segmented videos scored better on transfer tests, which measured the implementation of the learned material.

Other studies have also shown similar results. At Mayer, Dow and Mayer, learners were shown an animation of how an electric motor works [23]. Here the segmented version was linked with interactive elements. When the learner clicked on a question, the corresponding segment was shown. In the non-segmented version, a continuous video with identical content was shown. Again, the group with the segmented version performed better [23].

Boucheix and Guignard were also able to demonstrate the effect in slideshows [6]. The learners who received the segmented version again achieved measurably better results. They were able to influence the speed of learning themselves by clicking on the next-button, while two other groups were shown a fast and a slow, continuous version of the same slideshow.

In the technical literature, the pacing principle is regularly used synonymously with the segmenting principle. There is no desirable distinction between segmentation of learning videos with step control by the learners and segmentation with pauses given by the learning program without active control by the learners. Moreno and Mayer also use the term “pacing” as an alternative to segmentation [30]. In Doolittle, Bryant and Chittum segmentation is directly defined as control over pacing [11].

A segmentation without segment pauses forced by the program, but with chapter subdivision and short transitions, should in theory achieve the effect of increasing learning facilitation through better structuring of the learning material, since the learning material must also be structured in this type of segmentation and divided into logically connected individual sections. The second effect, the increase in learning facilitation by reducing the cognitive load, could be achieved at least partially if the transitions are designed in such a way that there is enough time for the learners to structure what they have already learned before the next chapter comes. Compared to segmentation with forced interruptions until the active interaction of the learning person (e.g. via a Next button), the pauses here are fixed by the video itself or the learning program, so that no adaptation to the level of knowledge or the processing speed of the learning person can take place.

A segmentation without pauses forced by the program was investigated by Spanjers et al. [42]. In the study they conducted, there were four conditions; not segmented, segmentation only via the brief blackening of the screen, segmentation only with pause, segmentation with pause and brief blackening of the screen.

A meta-analysis by Mayer and Pilegard [27] showed that a segmentation effect could be demonstrated in 10 of 10 studies examined. The median effect size was 0.79 which is a strong indicator that segmentation has a positive effect on learning.

## 5 EMPIRICAL STUDIES

On the basis of the available theoretical and empirical findings presented in the last section, there are still a number of open questions for the formulation of a design pattern:

- What is the recommended video length for different video types, production styles, learning scenarios?
- In how many parts should a longer video be divided?
- How long should a segment be?
- What features requires a video player to support the use of segmented videos?
- Does the design of the video player for segmented videos have an influence on the learning success?

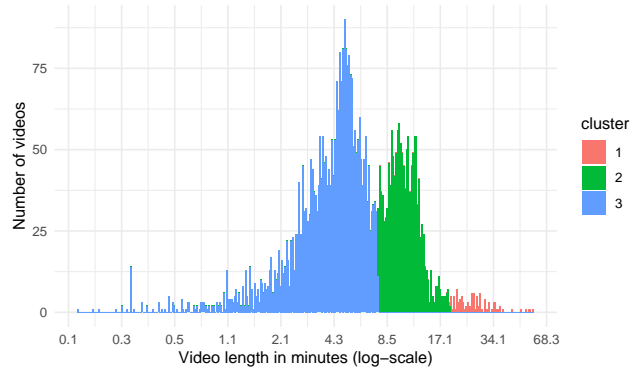
With regard to the creation of a design pattern on how to segment learning videos, we would need to answer these questions using empirical methods.

As a first step, we are interested in current practice in the production and application of educational videos. For this purpose we are studying a large number of relevant educational videos on YouTube. YouTube provides a great variety of video types and production styles. The naming of the videos also gives information whether the video is part of a series of a certain size. The number and duration of segments can thus be determined. However, statements about certain learning scenarios and the design of video players cannot be made on the basis of a single video platform. For this reason, we have conducted a user study as a second empirical pillar, investigating alternatives for the design of video players supporting segmentation.

### 5.1 Analysis of the video length of popular learning videos on YouTube

As the largest video platform in the world, YouTube is particularly well suited to determine video trends and analyse viewer behaviour. The trend towards greater acceptance and demand for shorter learning videos should therefore also be reflected here. The aim of this analysis is to examine the most successful or popular knowledge-sharing YouTube channels in terms of the length of their videos. For this purpose, a list of channels to be investigated must first be compiled and then the length of all videos must be documented.

**5.1.1 Method.** First, we found out which YouTube channels belong to the education sector. The question was asked once for "most popular" (list A) and once for "best" (list B) YouTube channels. The first 10 results of both search queries were taken into account. Since there were seven overlaps with the two search terms (see table), this resulted in a list of 13 websites. Each website has its own list of recommended channels. The number varies between 8 and 100, with the first 20 recommendations being taken into account. In the list "medium" no order could be found and all entries were considered. All channels that were recommended in more than 3 of the 13 lists, i.e. were part of the entries considered in more than 3 lists, were then taken into account. This was the case for the following 12 channels. The number of lists in



**Figure 2: Overall distribution of the video length clustered for video length**

which each channel was represented is shown in the second column of Table 3.

After that, we determined the popularity of the channels by ranking the number of page views. The channels to be evaluated were analyzed with the help of the web application "YouTube Playlist Analyzer". The application makes it possible to list the last 600 videos and display their length. Eight of the twelve channels in the list have less than 600 videos, so all existing videos are evaluated here. Of the four channels with more than 600 videos, the last 600 videos were considered.

To eliminate outliers, all videos that are 10 times longer than the average (e.g. hours long live broadcast) were excluded.

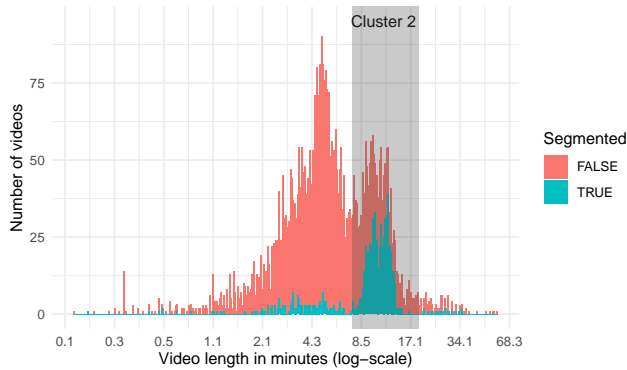
**5.1.2 Results.** The observation included 4,263 videos. The video length ranges from 9 to 3,417 seconds. The overall shape of the distribution of the video length is skewed to the left (see Fig. 2). The histogram indicates a multimodal distribution of at least two independent groups of videos.

In order to identify and describe possible groups we used K-Means as a clustering algorithm. The number of cluster was computed using the average silhouette of observations for  $k$  clusters with  $k = [1, \dots, 15]$ . The maximum of the average silhouette over the range values for  $k$  cluster was 3. Tab. 1 provides an overview of the clusters found in the data set. The smallest cluster contains 99 longer videos with a duration ranging from 19:25 to 56:55 minutes ( $M=27.70$ ,  $SD=7.84$ ). The cluster of videos with a medium length contains 1422 videos with a total length between 7:32 and 19:15 minutes ( $M=10.85$ ,  $SD=2.32$ ). The biggest cluster is the one of short videos. The duration of 3102 videos is between 9 seconds and 7:31 minutes ( $M=4.20$ ,  $SD=1.64$ ).

The remarkably large number of videos in clusters 2 and 3 raises the question whether these videos stand alone or are part of a series. Series can be understood as a continuation of a story or as a segmentation of a longer video. However, the difference between the two variants cannot be determined due

**Table 1: Clusters C for video length in seconds**

C	Videos		Min	Max	M	SD
	total	segm.				
1	99	19	1165.00	3417.00	1662.23	470.76
2	1422	570	452.00	1155.00	650.92	139.31
3	3102	156	9.00	451.00	251.92	98.18

**Figure 3: Overall distribution of video length of segmented and unsegmented videos**

to the missing date of release that we could not determine with the used tools.

*Regular Expressions* were used to identify videos from the sample whose titles contained indications of segmentation. Specifically, the keywords *part*, *episode*, *volume*, *period*, *#* – including plural forms and common abbreviations – were searched for in conjunction with digits. A manual review of the 745 extracted video titles did not find any false positives. As shown in Fig. 3, there are 570 segmented videos (76.5%) in cluster 2, 156 (20.9%) in cluster 3, and only 19 (2.6%) longer videos as part of a sequence in cluster 3. Thus, the YouTube channels considered in this study tend to produce and publish videos of medium and short length as part of a series of videos. The share of segmented videos in cluster 2 is 40.0%, in cluster 1 5.0%, and in cluster 3 at least 19.2%.

Apart from the overall distribution of the video length we have examined the individual channels in detail. As shown in Tab. 2 there is only one channel whose videos are on average longer than 10 minutes. The most common time frame is 3-5 minutes. Four channels fit into this range. It is noticeable that there are no channels whose videos are on average shorter than 3 minutes. This is also due to the fact that even channels that have a lot of videos under 3 minutes, such as "MinutePhysics", also have some longer videos that pull up the cut. The "typical" videos of these channels are thus each below the average running time indicated here.

**5.1.3 Discussion.** The selection of YouTube channel starts with a Google search for hit lists of popular and widespread educational videos. However, the search with Google is highly

personalized and was therefore not free from the subjective imprints of the experiment leader. The identified hit lists of YouTube channels are also not free of imprints. No detailed investigation was carried out into how these lists came about. Commercial intentions, thematic preferences or individual interests of the list makers may have influenced the compilation. This kind of disturbing systematic influences is also evident in the fact that, with a few exceptions, only English-language videos were included in the channels. Any cultural differences could therefore not be taken into account.

The multi-stage selection process was subject to several practical and technical compromises. Starting with the search for hit lists, the selection of lists and channels and the inclusion of 600 videos per channel, several decisions were made which could not be justified empirically. The sample of the learning videos available on YouTube that we are looking at here was therefore only selected quasi-randomly. The sample of  $N=4.263$  videos was also relatively small compared to the total number of learning videos. The number of 12 out of 841,317 YouTube channels is also very small, even if compared to the number of channels in the category *education*.

For the individual videos it could not be determined beyond doubt whether they were all learning videos in the strict sense. The content of the videos (e.g. topic, field), the production style (e.g. office video, presentation, khan-style), type of learning video (e.g. instruction, explanation, observation) and the target audience (e.g. K12, university students) were not examined in detail. Possible differences in video duration with respect to these dimensions were not investigated.

Each video was viewed individually. The context of videos which belong to a series or sequence of videos was determined by common keywords appearing in the video title. Semantic relations that were not expressed in the video title, but in the video itself could not be considered.

A technical feature of the video was used to determine the video length. The duration of the didactic intervention conveyed in the video could not be recorded. From the video length one would have to subtract as the time used for the opening credits, closing credits or other interruptions.

**5.1.4 Design implications.** The investigation of video length based on the length of video in popular educational YouTube channels has shown that there are three groups of videos of different length. The largest group two thirds of the videos which are shorter than 7:52 minutes. This result shows that the production of short videos has proven itself, at least within the successful YouTube channels. The production of videos with up to 8 minutes playing time can be considered as a common practice. However, videos with an length between 7:52 and 19:25 also make up a significant proportion of 30.8%, while longer videos are the exception.

With regard to the segmentation of educational videos, the channels considered here show a clear picture. 40.0% of the videos with medium length and 19.2% of the short videos are part of a series of segmented videos. This result shows that more comprehensive topics can be transported in the form of a video if they are divided into smaller parts. In the

**Table 2: Overview of the analysed channels.**

Channel name	Number of videos			Views in mio.		Video length (mm:ss)	
	total	sample	ratio	total	per video	M	SD
CrashCourse	1100	600	54.5	1035,3	0,94	10:31	02:26
Veritasium	265	265	100.0	509,9	1,92	05:30	03:08
Vsauce	378	378	100.0	1584,5	4,19	07:32	07:07
ASAP Science	314	314	100.0	1217,7	3,88	03:46	02:21
SciShow	1912	600	31.4	1003,3	0,52	07:30	06:08
MinutePhysics	232	232	100.0	395,4	1,70	03:05	02:36
TED-ED	1574	600	38.8	1321,7	0,84	04:44	00:59
CGP Grey	130	130	100.0	428,0	3,29	04:39	03:43
SmarterEveryDay	298	298	100.0	559,4	1,88	05:56	04:34
Khan Academy	7284	600	8.2	1658,4	0,23	06:12	04:51
Numberphile	518	518	100.0	398,5	0,77	08:55	05:16
Kurzgesagt	96	96	100.0	597,5	6,22	06:53	02:26
All	14101	4631	32.8	892,5	2,20	06:16	03:48

YouTube channels considered here, it is a common practice to divide videos into medium and short length segments.

However, the implications set out above apply only to the possible uses within YouTube. The videos of the channels examined here were produced exclusively for YouTube and the tools available there. Other video players may support other tools and therefore enable other usage patterns. The following study examines how video segmentation can be supported by video player design.

## 5.2 User study on segmentation of educational videos

The analysis of popular YouTube channels has revealed a certain practice by which videos are produced and used. However, the perspective of the users as well as other design options have not yet been considered. The goal of the following user study is to compare design variants for video players, each of which supports the segmentation of longer videos in different ways. The comparison of different player variants should lead to findings for the design of video players that are conducive to learning.

**5.2.1 Design of the study.** The study is a comparative study. Participants are randomly exposed to different treatments. The treatments consist of two video types and three segmentation variants – two treatments and one control condition. The test persons receive a short e-mail with explanations about the scope of the data collected and the content of the study. The attached link takes them to the web application where they can participate in the study. All participating persons go through the same process steps. First, some personal

data is collected in step 1. In step 2, previous knowledge of the topic dealt with in the video is asked for. In step 3 the participating persons are shown approx. 45 minutes of video material. In step 4 an evaluation of the shown learning material is asked for and in step 5 the same knowledge query is made as in step 2. After submitting the knowledge query again, participation is finished. It took about 5 to 10 minutes each to answer the two questionnaires at the beginning and at the end of the test participation. Together with a running time of the videos of about 45 minutes each, this results in a total time of about one hour for the test participants. In the end, the learning success is calculated and evaluated from the answers to the two knowledge questions.

**5.2.2 Material.** The study is basically two studies that were conducted with the same framework conditions for two different types of video. One is a lecture, which strongly resembles a lecture recording in structure and content, and the other is a TV documentary, similar to the TerraX format, which is broadcast regularly on Sunday on ZDF.

In the control condition (variant 1), the participants are allowed to watch a non-segmented video. The video is not divided into subchapters and does not contain chapter titles. Participants can pause and resume the video at any time or jump forward or backward as desired by clicking on the progress bar. The UI of the player as shown in Fig. 4 is therefore hardly different from a normal video player.

For variants 2 and 3 the video was segmented at previously selected points. Chapter transition animations were integrated at these points. The screen goes dark for 1-2 seconds and the title of the current segment is displayed. However,

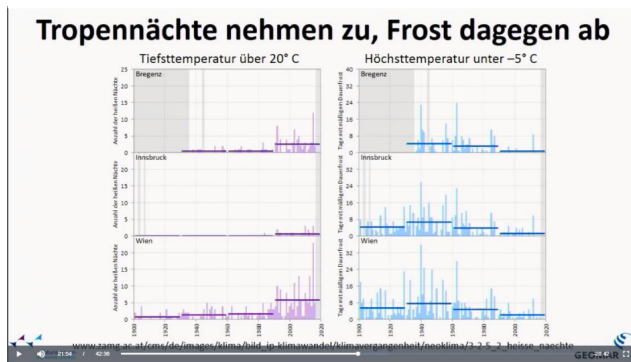


Figure 4: The video in variant 1 was not segmented. The video player provided only basic functionality.

both variants differ in the way the segmented content is presented. The learning environment of variant 2 (see Fig. 5) has been enhanced compared to variant 1 with a chapter overview, chapter boundaries and navigation options. In variant 2, the video learning environment has been changed so that a chapter overview is permanently displayed on the left side of the video player. There the test persons can view the chapter structure and jump directly to it by clicking on the chapter. The chapters are also displayed on the progress bar. Red markers distinguish the different chapters from each other in colour. There is also another button in the video player. Besides the progress bar, another chapter navigation option has been integrated. Thus, even in full screen mode, in which the left chapter overview would no longer be visible, it is possible to jump to any desired segment.

In variant 3, the video of variant 2, extended with chapter transition animations, was cut at the segment boundaries and individual videos were generated from it. The individual videos were not shown continuously, but were followed by a pause forced by the learning environment, which could only be ended by the interaction of the test participants with the learning environment. As can be seen in Fig. 6, the individual videos with thumbnails are displayed above the video player. By clicking on the images or the titles directly below, the corresponding video is loaded. The next video is loaded by clicking on one of the thumbnails or on the button "Continue to chapter [x]". The participating person has to actively decide whether the next video should follow directly, or whether the person wants to reflect briefly on what he has learned or perhaps watch some content again. In addition, the title of the currently shown video is displayed directly above the player.

**5.2.3 Collected data.** Since the study was conducted online and the investigators had no personal contact with the participants, socio-demographic data, knowledge acquisition and feedback on the respective intervention were collected by means of four questionnaires and video analytics methods.

- Personal data: In detail, demographic data (age, gender, educational level), video consumption behaviour,

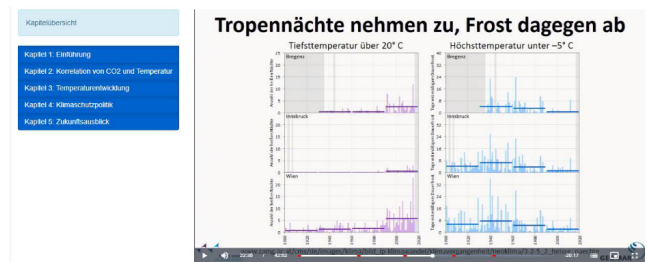


Figure 5: Variant 2 with segmented chapters which were labeled on the timeline and in a table of contents

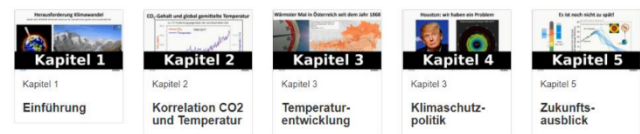


Figure 6: Variant 3 splitted chapters into separate video vignets. The chapter videos were provided as in variant 1, but with a thumbnail menu (see above) linked to remaining chapter videos.

interest in the topic of the video and a self-assessment of learning and memory skills were collected. Questionnaire

- *Pre-test:* A test of the participants' prior knowledge, adapted to the video. The test consisted of 10 multiple-choice questions.
- *Video analytics:* During the video playback clickstream data and timeupdate events are collected in a logfile.
- *Feedback:* Participants evaluate the learning environment in terms of entertainment, structure, information content and length. Additionally, questions about the learning content, climate change, are asked again.
- *Post-test:* Repeat performance of the pre-test. The learning success results from the difference between the score of pre-test and post-test.

**5.2.4 Participants.** Computer science students and colleagues were invited to participate in the survey by e-mail. It is estimated that 100 to 150 people received the invitation, since the invitation mail has been forwarded to instant messaging channels among others. Of these, 38 people called up the web application and filled out the first questionnaire (20 for the lecture video and 18 for the documentary). During the second step of the study (pre-test before watching the videos, 4 persons ended the experiment prematurely (all of them were assigned to the lecture video). Of the remaining 34 persons, another 2 persons interrupted their participation in the study while watching the video (one person per video type). Of the remaining 32 persons, 6 more persons briefly watched the video and then proceeded to the next step, only stopping the experiment there without editing the two questionnaires afterwards (2 were assigned to the lecture video, 4 were assigned to the documentary).

**Table 3: Distribution of participants per treatment and video type**

Type of video	Variant 1	Variant 2	Variant 3	sum
Documentary	4	4	4	12
Video lecture	3	3	4	10
sum	7	7	8	22

Of the 26 persons who filled in all questionnaires, 4 were not considered because they had seen less than 10% of the video material (all of them were of the lecture type). In total, 22 of the 38 persons who had started the study were still able to evaluate the data. This corresponds to a rate of 57.8%. In the following, the number of participants will therefore always refer to the 22 persons (14 male, 8 female). 17 participants, and thus the majority, were between 25 and 45 years old. 2 participants were under 25 years of age and 3 participants were over 45 years old. Table 3 shows how the participants are divided between the three treatments and the two video types.

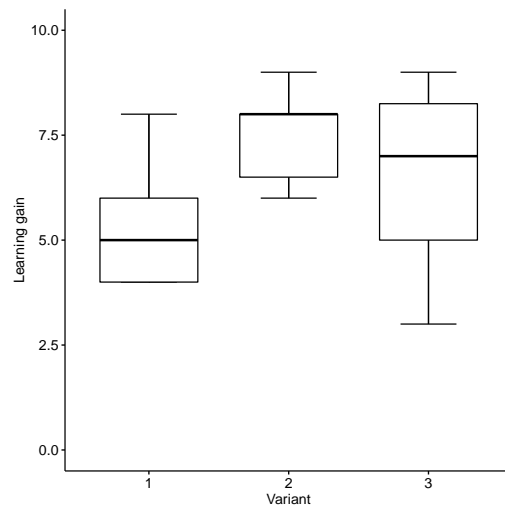
**5.2.5 Hypotheses.** The following hypotheses have been tested:

- H1 Segmented videos with sequences of up to 10 minutes in length (variant 3) achieve a higher learning effect than long, non-segmented videos (variant 1).
- H2 Segmented videos (variants 2 and 3) are perceived by viewers as better structured than non-segmented videos (variant 1).
- H3 Long videos are just as conducive to learning as segmented individual videos (variant 3), provided that the video learning experience is supported by a TABLE OF CONTENTS or chapter index, navigation options, short chapter transitions and visible chapter boundaries (as markers on the ANNOTATED TIMELINE) (variant 2).

**5.2.6 Implementation.** Since the invitation e-mail contained a link to the web application the recipients were able to start the app immediately. When opening the app the first time the participants were randomly assigned to a video type (lecture or documentary) and segmentation variant. During and after the study the investigators were not communicating with the participants. The conduct of the study was not actively monitored by investigators. There were no problems during implementation or requests for assistance. The video usage data was collected by the web server and the survey results were stored in a Google spreadsheet.

### 5.2.7 Results.

**Learning gain.** For all videos and across all groups, participants had minimal prior knowledge (see Tab. 4). The topics covered in the documentary were intellectually less demanding, which is why the participants scored slightly better in the pre-test compared to the lecture. The participants assigned to the three variants hardly differed from one another in terms of their prior knowledge. Only the participants in the control group (variant 1) recorded slightly better results ( $M=1.29$ ,  $SD=1.11$ ).

**Figure 7: Boxplot of learning gains by variant**

The post-test showed larger differences in performance. The differences between the two video types are only marginal, but the learners with variant 2 and 3 achieved significantly better results than the control group with variant 1. Compared to the control group, the learners with variant 2 achieved a good 22% better test results ( $M=8.43$ ,  $SD=0.98$ ). With variant 3 the learners achieved an increase of 13% ( $M=7.60$ ,  $SD=1.85$ ) compared to the control group ( $M=6.57$ ,  $SD=1.62$ ).

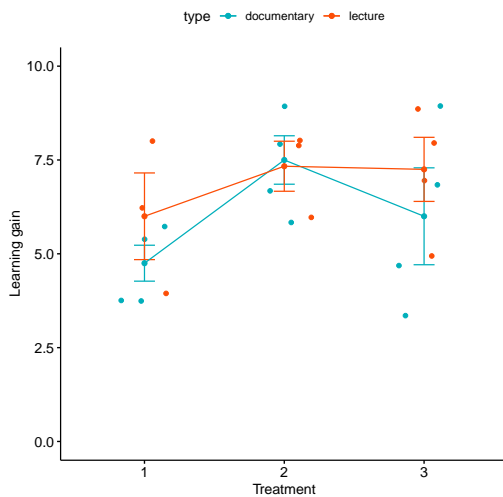
The differences in post-test are reflected in the learning gain. Compared to the control group, the participants scored almost 29% more points with variant 2. And also with variant 3 the results were better by 22%. Fig. 7 and 8 show the distribution of leaning gains for each group.

An analysis of variance (ANOVA) with the variant 1 and 2 as a between-subjects factor and learning gain as a within-subjects factor revealed significant for learning gain,  $F(1, 14) = 9.12$ ,  $MSE = 1.7619$ ,  $p < .01$ . ANOVA with the variant 1 and 3 as a between-subjects factor and learning gain as a within-subjects factor resulted no significant effect for learning gain,  $F(1, 15) = 1.9216$ ,  $MSE = 3.49$ ,  $p = ns$ . Consequently, H1 can be accepted.

**Video interactions.** Based on the evaluations of other studies (cf. [4, 33]), it was to be expected that the pause function would be used only rarely. These expectations were confirmed. Only three participants used the pause function once each. Since there was no field study here, it can be assumed that the participants wanted to complete those presented in the study quickly and without interruptions. The paused periods were 27 seconds (type documentation variant 1), 89 seconds (type presentation variant 2) and 5 seconds (type presentation variant 1). In all cases these are participants of a group who were shown a variant without automated breaks. Participants in group 3, who were shown the individual videos segmented with forced pauses, did not pause, but on average allowed more time to pass before clicking on to the next video.

**Table 4: Mean pre-test and post-test scores and learning gains for each group and type of video.**

Type of video	Variant	n	Pre-test		Post-test		Gain	
			M	SD	M	SD	M	SD
Documentary	1	4	2.00	0.81	6.75	0.96	4.75	0.96
Documentary	2	4	1.25	1.89	8.75	0.96	7.50	1.29
Documentary	3	4	1.50	0.58	7.50	2.08	6.00	2.58
Documentary	1-3	12	1.58	1.17	7.67	1.56	6.08	1.98
Video lecture	1	3	0.33	0.58	6.33	2.51	6.00	2.00
Video lecture	2	3	0.67	0.58	8.00	1.00	7.33	1.15
Video lecture	3	4	0.50	1.00	7.75	1.89	7.25	1.71
Video lecture	1-3	10	0.50	0.71	7.40	1.84	6.90	1.60
All	1	7	1.29	1.11	6.57	1.62	5.29	1.49
All	2	7	1.00	1.41	8.43	0.98	7.43	1.13
All	3	8	1.00	0.93	7.60	1.85	6.63	2.13
All	1-3	22	1.09	1.11	7.55	1.65	6.45	1.82

**Figure 8: Distribution of learning gains by variant**

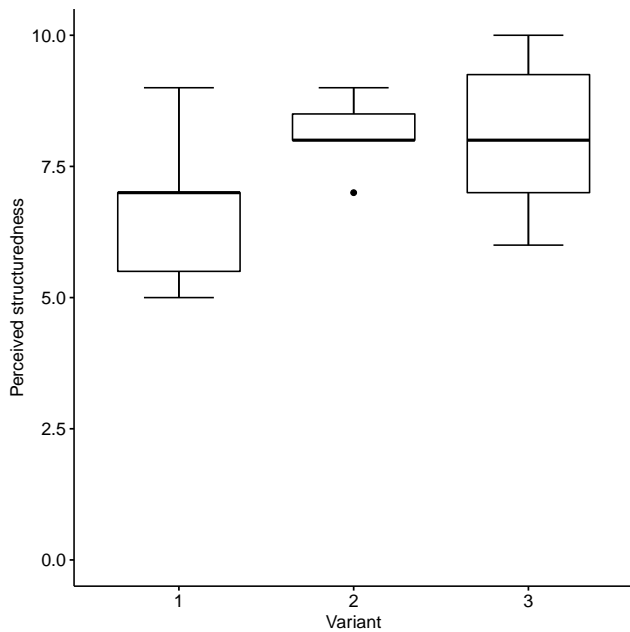
Usually 30 seconds to 2 minutes. One person waited over 3 hours between two video segments. It is conceivable that another activity was started during this time and the test participation continued afterwards.

Compared to the breaks, jumps in the video were used relatively often. 14 of 22 participants used the possibility to jump forward or backward on the progress bar in the video. The participants who used this possibility jumped on average 3 times while watching the video in one or the other direction. Of the 42 recorded jumps, 34 were forward and 8 backward. When jumping backwards, it was mostly only a few seconds (approx. 7 to 15 seconds), when jumping forward, however, it was longer distances (approx. 30 seconds to 4 minutes). There was no repeated playing of a chapter or a longer passage, the few rebounds that were performed were usually only a few seconds. It is conceivable that something was overheard here which should be played again. It is also conceivable that the test persons remembered the questions of the prior knowledge

test and recognized that what they had just heard or seen contained an answer to a question and should be looked at again. There were also strong outliers in the jumps. One participant jumped forward by 36 minutes directly after the second second. One explanation would be that there was a technical problem (e.g. browser crashed) and the participant had to reload the page.

Participants of group 1 were provided the non-segmented version with the standard player, accordingly they could not select chapters or segments. Participants of group 2 were able to orientate themselves by the chapter markers on the progress bar and to jump to a chapter at will via the chapter menu. Out of 8 people in group 2, only 2 people did not use the chapter menu and simply let the video scroll through from start to finish. Two people used the table of contents to jump to a desired location only once, three people three times, and one person four times. Group 3 participants had to click either on the "Next" button or on one of the chapter thumbnails to move on to the next video. All participants followed the given order of the chapters. There was no non-sequential viewing of the videos. If chapters or sections were left out, they were not viewed later.

*Perceived structuredness.* The participants were asked to state how well structured they perceived the video material shown. Both segmentation variants were rated on average 23% as better structured than the non-segmented video (see Fig. 9). For both types of video, a clear picture emerges with regard to the evaluation of the structuring. The advantages of segmentation on learning success identified in other studies are partly attributed to the increase in learning facilitation through better structuring of the learning material. The data collected here confirm that the improvement in structuring is also perceived and confirmed by the viewers. It is striking that the evaluation of structuring, more than all other variables discussed here, hardly differs according to video type. The data suggest that there is a clearly discernible tendency to perceive segmented videos as more structured if they have



**Figure 9: Perceived structuredness for each variant**

short chapter transitions and chapter titles. Consequently, H2 can be accepted.

*Segmentation support.* Variants 2 and 3 supported the segmentation of longer learning videos in different ways. The extent to which both variants can be regarded as conducive to learning should be determined in this study.

The comparison between variant 3 and 2 provides a mixed picture. With regard to the perceived structuring, both still achieved similarly good results. Variant 2, however, received an overall rating 6.8% better than variant 3 and a 10.5% better rating of the information content, but was rated 5% worse in the learners' self-assessment, has a lower probability of being recommended to others and was also rated more often and more strongly than too long.

With regard to the measured learning success relevant to the hypothesis, variant 2 ( $M = 7.25$ ,  $SD = 1.16$ ) scored 9.4% better than variant 3 ( $M = 6.63$ ,  $SD = 2.13$ ). The chapter menu and the visible chapter boundaries were able to raise this segmentation variant to the same high level as the variant segmented with individual videos with regard to structuring. The chapter menu not only improved the structure and served as an identifier for the selected chapters, but was also actively used as an interactive navigation element by the test persons. As the evaluation of the interactions with the video player showed, these possibilities were used by 8 of the 10 participants of group 2.

It is noticeable that the given sequence of the segments, which were numbered with a chapter number, was adhered to by all test persons. The navigation options were thus used to jump from the current segment to the beginning of the

next segment or to skip a segment completely - but not to view segments in any order.

### 5.3 Discussion

The number of 22 participants was relatively small. Due to the examination of longer videos, the participants had to invest a considerable amount of time without receiving any financial or non-material incentives. An expense allowance might have increased the motivation to participate in the study and reduced the drop out.

Since the study was conducted online, the participants could not be monitored under laboratory conditions. We assume, however, that the participants performed all tasks independently and without outside help or other aids.

Regarding the video reception, it was not possible to determine how much attention the participants paid to the learning content. At least one participant in the experiment proved that he let almost 2 hours pass before continuing with the tasks he was given after watching the video. For technical reasons it was unfortunately not possible to record the exact playback behaviour. A connection between the learning effect and the playback behaviour (e.g. multiple viewing of sections) could not be determined.

The videos examined here both dealt with climate change. It is possible that other domain-specific topics with the player variants examined here lead to better or worse learning outcomes. The same applies to effects due to production styles other than lecture and documentation. Neither were target group-specific effects (e.g. K12, university students, vocational training) considered. Finally, it cannot be ruled out that certain learning scenarios may favour or restrict learning with one of the player variants.

The player variants were firmly defined. In variants 2 and 3, several design elements were combined. In detail, these design elements could be differentiated more precisely and presented in a different way that might be more conducive to learning.

The experimental groups were not pre-selected according to their level of knowledge. Although the level of prior knowledge was queried and can be compared among the test participants, the 10 multiple-choice questions are hardly suitable for grouping dedicated knowledge level differences. Disturbing variables such as the handling of ignorance can falsify the results. Some participants chose "Don't know" for questions whose answer they did not know, others chose any answer in case of ignorance. In addition to the "Don't know" option, there were usually five other possible answers. With five answers, the strategy of randomly choosing between the possible answers would provide the correct answer on average in 20% of cases. With only 10 questions, this can skew the evaluation of the participants' prior knowledge level. In addition, the learning content shown in both video types was aimed at people with no prior knowledge about climate change. Accordingly, the content is designed in such a way that all learners are assumed to have no prior knowledge of

the topic. Experts could probably only benefit from their previous knowledge to a very limited extent.

The multiple-choice questionnaires used were mainly used to measure the participants' memory and attention immediately after watching the video. Statements about the long-term learning effect require the implementation of follow-up tests.

Finally, the results cannot be generalised to draw conclusions about learners who learn with videos because the sample was not random and could be biased. So the results may not be representative for a population of video learners.

#### 5.4 Design implications

The study has shown that video segmentation should be supported by the design of the video player. The player used in variant 2 differs significantly from standard players used by the control group with variant 1. In addition to a TABLE OF CONTENTS, a ANNOTATED TIMELINE with the segment boundaries, temporal pauses with identification of segment transitions and titles were realized.

Also the division of a long video into shorter segments led to better learning results if the individual segments were represented in a sequence to be called up individually. YouTube, for example, offers the possibility to list related videos within PLAYLISTS for the video segments identified in the YouTube EDU channels.

From a technical point of view, it is remarkable that the design options presented in variants 2 and 3 can be realized completely in the player without splitting the underlying longer video.

### 6 DESIGN PATTERN

The procedure for the empirical elaboration of design patterns presented in section 3 started with the pattern SIX MINUTE VIDEO by Mor and Warburton [28]. The hypothesis expressed therein consists of three statements:

- A "create smaller discrete episodes" (six minute video)
- B "episodes that you can feel confident will be watched in their entirety by your learners."
- C "Ensure that the material is designed within a structure for REGULAR ATTENTION RECUPERATION<sup>2</sup>

In the first study (see section 5.1) statement A was partially confirmed. YouTube educational videos are usually shorter than 8 minutes, or shorter than 20 minutes if they are part of a series. So there is some evidence that the production of shorter episodes is a common practice. The guideline value of "6 minutes" is in line with the average length of all videos viewed, but without taking into account the distribution of video length. The assumption made in the pattern of Mor

<sup>2</sup>REGULAR ATTENTION RECUPERATION: "Include delivery forms and activities in your lecture that help with regaining the attention of the students. Vary in the sorts of activities/delivery forms and involve the students actively in these attention refreshments". Mor and Warburton [28]."

and Warburton [28] is thus consistent, but not differentiated enough.

Regarding statement B, no statements on the duration and intensity of the video observation could be made in the second study (see section 5.2). On the basis of the learning gain, however, it can be concluded that the videos were watched almost completely. Thus, it could be shown that segmentation is more conducive to learning, which supports statements A and B.

Statement C is unfortunately very imprecise. However, there is an indirect demand to interrupt the video viewing by forms of presentation and built-in activities and thus attract the attention of the learners. In the second study we purposed the variants 2 and 3 to investigate suitable designs for video players that meet these requirements. Statement 3 could thus be made more precise with regard to concrete design.

In view of this critique of the pattern SIX MINUTE VIDEO, which are only exemplary described here, a new pattern had to be written, which takes into account the empirical results. The design pattern called VIDEO SEGMENTATION presented below is limited to the types of videos investigated, includes literature findings and mentions our own empirical studies.

#### Video Segmentation

**Divide longer videos into meaningful segments of up to 20 minutes and support the transition between the segments in the player.**

Educational videos such as lecture recordings, documentaries, and explanation videos.



**Learners do not watch longer learning videos until the end and thus miss relevant learning content.**

- *Cognitive load*: The constant flow of information generates a high cognitive load on the viewer if the information follows each other too quickly [43] like it is often the case in a video.
- *Production effort*: Producing a video is much effort compared to text or audio representations of the same content.
- *Dramaturgy*: The dramaturgy of a learning video or the rhetorical structure of a lecture is not accidental, but intended by the presenter. This dramaturgical structure cannot be broken up easily.
- *Breaks*: If someone interrupts the playback in the course of the video, there is a chance to continue watching at that point later.
- *Information seeking*: If you are specifically looking for information in the video and have found it, it is a legitimate reason to stop the playback and turn to other activities.



**Divide longer videos into meaningful segments of up to 20 minutes. Mark the boundaries of the segments clearly. Present the segments as a labeled sequence of related videos.**

The segments must be defined manually. It is advisable to use existing segments, such as section headings or content parts. Segments should never be defined on the basis of rigid temporal intervals, but should comprise semantic units. Advances in the detection of sections and paragraphs using natural language processing methods based on video TRANSCRIPTS have not yet been used to create meaningful segments.

The length of a segment should be less than 20 minutes. A longer video should be divisible into at least two segments. If there are more than 10 segments, it will be difficult for the learner to keep track of them. In this case an additional hierarchy level should be used. As a result typical lecture of 90 minutes length can be segmented into five to six pieces.

From a technical perspective a segmentation can be performed in different ways within a video player without changing the source material of the video. The beginning of a segment can be marked by a still image or blank screen showing the corresponding title of the segment. If the video does not pause automatically, the viewer should be given enough time to adjust to the coming segment. Breaks have a double function. On the one hand, they give the learner time to recover and recapitulate, on the other hand, they emphasize the end of a semantic unit. These units can also be presented in the form of a TABLE OF CONTENT or as a list of SEQUENTIAL MEDIA elements to facilitate precise jumps between segments. In order to recognize not only the topic of a segment by the labels displayed, but also its temporal extension, the segments should be annotated on the timeline (cf. ANNOTATED TIMELINE).

### Benefits.

- *Cognitive load*: Both, structuring of the learning material and the learning breaks made possible by the segmentation can help the learner to reduce the cognitive load.
- *Production effort*: A subsequent editing and splitting of the video material is not necessary. Existing videos can be re-used if the segments have been marked and labeled.
- *Dramaturgy*: The intended structure of a video becomes transparent by the TABLE OF CONTENT and the labeled segment boundaries on the timeline (cf. [41]).
- *Breaks*: Breaks help the learner to conclude with a unit of content. Breaks are necessary to reactivate cognitive resources.
- *Information seeking*: Named segments can help learners get an overview of video content and quickly find the information they are looking for. For a more detailed search, however, further technical aids such as full-text search in a TRANSCRIPT, TEMPORAL TAGS, or TEMPORAL BOOKMARKS may be required.

### Liabilities.

- *Cognitive load*: The proposed solution doesn't force the learner to make a break at the end of a segment.
- *Production effort*: Segments need to be annotated for existing video footage. Segment boundary may also be generated by the learners [36]. However, this is a very time-consuming and cognitively demanding task.
- *Dramaturgy*: There is a risk that segments will be torn out of context as a result of the division. The fact that the segments build up on each other enabled previous knowledge to be built up and assumed in a later segment. However, if learners do not look at the previous segments, they may lack previous knowledge and understanding of the current segment.
- *Breaks*: Breaks may disrupt the flow of learning if they occur in unfavorable positions.
- *Information seeking*: Named segments can help learners get an overview of video content and find the information they are looking for more quickly. For a more detailed search, however, further technical aids such as full-text search in a TRANSCRIPT, TEMPORAL TAGS, or TEMPORAL BOOKMARKS may be helpful.

**Empirical background.** In an analysis of 10 popular YouTube Channels 745 segmented videos (18.0%) were found in a sample of 4.136 video. 76.5% of the segmented videos were between 7:52 and 19:25 minutes long, 20,9% of the segmented videos were shorter than 7:52 minutes.

In a user study (N=22) segmented videos with sequences of up to 10 minutes resulted in higher learning effects than the non-segmented version of the same video.

### Known uses.



Figure 10: At *Coursera* MOOCs are split up into several sequences of short videos.

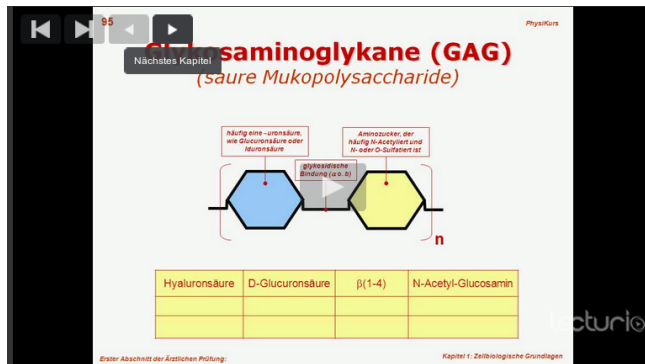
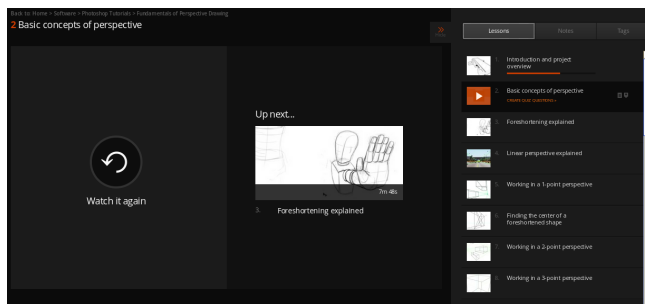
Figure 11: *Lecturio*Figure 12: *Digital tutors*

Figure 13: *iVersity* was a former MOOC provider whose courses consisted of several units, which in turn consisted of short videos.

#### Related Patterns.

In the drafted pattern SIX MINUTE VIDEO Mor and Warburton [28] emphasized the need for limiting the video length, but the segmentation of longer videos has not been addressed.

The patterns TABLE OF CONTENT, SEQUENTIAL MEDIA, and PLAYLIST as well as BASIC CONTROLS and ANNOTATED TIMELINE introduced by Seidel [40] complement the segmentation of learning videos.

## 7 CONCLUSION

While pattern mining is characterized by empirical examinations the applied solutions propagated in the patterns are only rarely the subject or the result of empirical tests. However, confidence that a pattern will serve its purpose can only be obtained through tests, evaluations and their application in the field.

This article uses the pattern VIDEO SEGMENTATION as a use case to present an iterative approach on how to integrate results of empirical studies into the elaboration of design patterns. This approach aims to explore the design space and specifies the context of a pattern. For this purpose we analyzed educational YouTube channels regarding their structure and video length. In a user study (N=24) we compared two design variants with a control condition for two types of segmented videos.

Longer videos should be divided into meaningful segments of 6-8 minutes length. The segment boundaries should be marked clearly and titled using the video player. A combination with further patterns proved to be beneficial for the learners.

With the approach presented here, we want to encourage researchers from different disciplines to communicate their empirical research results in the form of design patterns.

## ACKNOWLEDGEMENTS

This paper would not have been possible without the data collected by Matthias Maaß in his master thesis. I am very grateful to him for the preliminary work. In addition, I want to thank my shepherd Michael Krisper for reviewing this paper and giving precious and helpful feedback. Also the feedback of all participants of the online writers' workshop at the *EuroPLoP 2020* conference was highly appreciated and therefore I want to thank all of them.

## REFERENCES

- [1] Mattias Arvola. 2006. Interaction Design Patterns for Computers in Sociable Use. *International Journal on Computer Application Technology* 25, 2/3 (2006), 128–139. <https://doi.org/10.1504/IJCAT.2006.009063>
- [2] Rune Baggetun, Ellen Rusman, and Caterina Poggi. 2004. Design Patterns For Collaborative Learning : From Practice To Theory And Back. In *World Conference on Educational Media and Technology, EDMEDIA'04*.
- [3] Antony W. Bates. 2015. *Teaching in a digital age – Guidelines for designing , teaching and learning*. open.bccampus.ca. k.A. pages. <http://opentextbc.ca/teachinginadigitalage/>
- [4] Nicolas Biard, Salomé Cojean, and Eric Jamet. 2018. Effects of segmentation and pacing on procedural learning by video. *Computers in Human Behavior* 89 (2018), 411–417. <https://doi.org/10.1016/j.chb.2017.12.002>
- [5] Staffan Björk, Sus Lundgren, Hugo Grauers, and Se Göteborg. 2003. Game Design Patterns. In *in Level Up: Digital Games Research Conference 2003*. Faculty of Arts, Utrecht University, Utrecht, 4–6.
- [6] Jean-Michel Boucheix and Hélène Guignard. 2005. What animated illustrations conditions can improve technical document comprehension in young students? Format, signaling and control of the presentation. *European Journal of Psychology of Education* 20, 4 (2005), 369–388. <https://doi.org/10.1007/BF03173563>
- [7] Cynthia J Brame. 2016. Effective Educational Videos: Principles and Guidelines for Maximizing Student Learning from Video Content. *CBE—Life Sciences Education* 15, 4 (2016), es6.

- <https://doi.org/10.1187/cbe.16-03-0125>
- [8] Nicholas Chen and Maurice Rabb. 2010. A pattern language for screencasting. In *Proceedings of the 16th Conference on Pattern Languages of Programs (PLOP '09)*. ACM, New York, 6:1–6:11. <https://doi.org/10.1145/1943226.1943234>
- [9] Ruth Colvin Clark, Frank Nguyen, and John Sweller. 2006. *Efficiency in Learning - Evidence-Based Guidelines to Manage Cognitive Load*. Pfeiffer, San Francisco.
- [10] James O Coplien. 1997. A pattern language for writers' workshops. In *Proceedings of the 2nd European Conference on Pattern Languages of Programming (EuroPLOP '97)*. Siemens Technical Report 120/SW1/FB. Siemens, Munich, Germany, 51–60. <http://dirkriehle.com/community-service/hillside-group/europlop-1997/p2final.pdf>
- [11] Peter E Doolittle, Lauren H Bryant, and Jessica R Chittum. 2015. Effects of degree of segmentation and learner disposition on multimedia learning. *British Journal of Educational Technology* 46, 6 (2015), 1333–1343. <https://doi.org/10.1111/bjjet.12203>
- [12] Dieter Euler. 2014. Design Research – a paradigm under development. In *Design-Based Research*, Dieter Euler and Peter Sloane (Eds.). Zeitschrift für Berufs- und Wirtschaftspädagogik, Vol. 1. Franz Steiner, Stuttgart, 15–44.
- [13] Dieter Euler. 2014. Design-Research – ein Paradigma unter development. In *Design-Based Research*, D. Euler and P.F.E. Sloane (Eds.). Franz Steiner Verlag, Stuttgart, 15–44.
- [14] Philip J Guo, Juho Kim, and Rob Rubin. 2014. How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos. In *Proceedings of the First ACM Conference on Learning @ Scale Conference (L@S '14)*. ACM, New York, NY, USA, 41–50. <https://doi.org/10.1145/2556325.2566239>
- [15] Neil Harrison. 1999. The Language of Shepherds: A Pattern Language for Shepherding. In *Proceedings of the 6th Annual Conference on the Pattern Languages of Programs (PLOP)*. Monticello, Illinois, USA. [http://jerry.cs.uiuc.edu/~sim\\$lop/lop99/proceedings/harrison/shepherding4.pdf](http://jerry.cs.uiuc.edu/~sim$lop/lop99/proceedings/harrison/shepherding4.pdf)
- [16] Alan R Hevner, Salvatore T March, Jinsoo Park, and Sudha Ram. 2004. Design Science in Information Systems Research. *MIS Q.* 28, 1 (2004), 75–105. <http://dl.acm.org/citation.cfm?id=2017212.2017217>
- [17] Christian Kohls. 2014. *The Theories of Design Patterns and their Practical Implications exemplified for E-Learning Patterns*. Dissertation. Katholische Universität Eichstätt-Ingolstadt.
- [18] Christian Kohls and Stefanie Panke. 2010. Is That True...?: Thoughts on the Epistemology of Patterns. In *Proceedings of the 16th Conference on Pattern Languages of Programs (PLOP '09)*. ACM, New York, NY, USA, 9:1–9:14. <https://doi.org/10.1145/1943226.1943237>
- [19] Jack Koumi. 2006. *Designing Video and Multimedia for open and flexible learning*. Routledge, Abingdon.
- [20] Christine Kunzmann, Andreas Schmidt, and Johanna Pirker. 2016. Pattern-oriented Approaches for Design-based Research in Collaborative Research Projects: A Knowledge Maturing Perspective. In *Proceedings of the 21st European Conference on Pattern Languages of Programs (EuroPLOP '16)*. ACM, New York, NY, USA, 19:1–19:8. <https://doi.org/10.1145/3011784.3011805>
- [21] Nan Li, Lukasz Kidziński, Patrick Jermann, and Pierre Dillenbourg. 2015. *MOOC Video Interaction Patterns: What Do They Tell Us?* Springer International Publishing, Cham, 197–210. [https://doi.org/10.1007/978-3-319-24258-3\\_15](https://doi.org/10.1007/978-3-319-24258-3_15)
- [22] David Martin, Tom Rodden, Mark Rouncefield, Ian Sommerville, and Stephen Vill. 2001. Finding Patterns in the Fieldwork. In *European Conference on Computer Supported Cooperative Work*, W Prinz, M Jarke, Y Rogers, K Schmidt, and V Wulf (Eds.). Kluwer Academic Press, Bonn, 39–58.
- [23] R Mayer, J Dow, and Susan E Mayer. 2003. Multimedia learning in an interactive self-explaining environment: What works in the design of agent.
- [24] Richard E. Mayer. 2009. *Multimedia Learning - second edition*. Cambridge University Press, New York.
- [25] Richard E Mayer and Paul Chandler. 2001. When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology* 93, 2 (2001), 390–397. <https://doi.org/10.1037/0022-0663.93.2.390>
- [26] Richard E. Mayer and Ruth Colvin Clark. 2012. *Applying the Segmenting and Pretraining Principles: Managing Complexity by Breaking a Lesson into Parts*. John Wiley & Sons, Ltd, Chapter 10, 204–220. <https://doi.org/10.1002/9781118255971.ch10>
- [27] Richard E Mayer and Celeste Pilegard. 2014. *Principles for Managing Essential Processing in Multimedia Learning: Segmenting, Pre-training, and Modality Principles* (2 ed.). Cambridge University Press, 316–344. <https://doi.org/10.1017/CBO9781139547369.016>
- [28] Yishay Mor and Steven Warburton. 2016. Patterns for using video in MOOCs. In *Proceedings of the 21st European Conference on Pattern Languages of Programs*. ACM, New York, NY, USA, [in Druck].
- [29] Roxana Moreno. 2007. Optimising learning from animations by minimising cognitive load: cognitive and affective consequences of signalling and segmentation methods. *Applied Cognitive Psychology* 21, 6 (2007), 765–781. <https://doi.org/10.1002/acp.1348>
- [30] Roxana Moreno and Richard Mayer. 2007. Interactive Multimodal Learning Environments. *Educational Psychology Review* 19, 3 (2007), 309–326. <https://doi.org/10.1007/s10648-007-9047-2>
- [31] Symeon Retalis, Petroskakis Georgia, and Yannis Dimitriadis. 2006. Eliciting Design Patterns for E-Learning Systems. *Computer Science Education* 16, 2 (2006), 105–118. <https://doi.org/10.1080/08993400600773323>
- [32] Evan F Risko, Dawn Buchanan, Srdan Medimorec, and Alan Kingstone. 2013. Everyday attention: Mind wandering and computer use during lectures. *Computers & Education* 68 (2013), 275–283. <https://doi.org/10.1016/j.compedu.2013.05.001>
- [33] S Guttormsen Schär and P G Zimmermann. 2007. Investigating Means to Reduce Cognitive Load from Animations. *Journal of Research on Technology in Education* 40, 1 (2007), 64–78. <https://doi.org/10.1080/15391523.2007.10782497>
- [34] Sandra Schön and Martin Ebner. 2013. *Gute Lernvideos ... so gelingen Web-Videos zum Lernen!* Books on Demand, Norderstedt.
- [35] Till Schümmer. 2005. *A Pattern Approach for End-User Centered Groupware Development*. Dissertation. FernUniversität Hagen.
- [36] Niels Seidel. 2013. Peer Assessment und Peer Annotation mit Hilfe eines videobasierten CSCL-Scripts. In *DeLFI 2013 – Die 11. e-Learning Fachtagung Informatik der Gesellschaft für Informatik*, Andreas Breiter and Christoph Rensing (Eds.). Gesellschaft für Informatik, Bonn, 83–94.
- [37] Niels Seidel. 2015. Making web video accessible – interaction design patterns for assistive video learning environments. In *Proceedings of the 20th European Conference on Pattern Languages of Programs*. ACM, New York, NY, USA, 17:1–17:16. <https://doi.org/10.1145/2855321.2855339>
- [38] Niels Seidel. 2017. Empirical Evaluation Methods for Pattern Languages: Sketches, Classification, and Network Analysis. In *Proceedings of the 22st European Conference on Pattern Languages of Programs*. ACM, New York, NY, USA, 13:1–13:24. <https://doi.org/10.1145/3147704.3147719>
- [39] Niels Seidel. 2018. Aufgabentypen für das Zusammenspiel von E-Assessment und Lernvideos. In *Videocampus Sachsen – Machbarkeitsuntersuchung*, Aline Bergert, Anke Lehmann, Maja Liebscher, and Jens Schulz (Eds.). Technische Universität Bergakademie Freiberg, 45–60.
- [40] Niels Seidel. 2018. *Interaction Design Patterns und CSCL-Scripts für Videolernumgebungen*. TU Dresden, Dresden. 380 pages. <http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa-233756>
- [41] Ingrid A E Spanjers, Tamara van Gog, and Jeroen J G van Merriënboer. 2010. A Theoretical Analysis of How Segmentation of Dynamic Visualizations Optimizes Students' Learning. *Educational Psychology Review* 22, 4 (2010), 411–423. <https://doi.org/10.1007/s10648-010-9135-6>
- [42] Ingrid A E Spanjers, Tamara van Gog, Pieter Wouters, and Jeroen J G van Merriënboer. 2012. Explaining the segmentation effect in learning from animations: The role of pausing and temporal cueing. *Computers & Education* 59, 2 (2012), 274–280. <https://doi.org/10.1016/j.compedu.2011.12.024>
- [43] John Sweller, Jeroen J G van Merriënboer, and Fred G W C Paas. 1998. Cognitive Architecture and Instructional Design. *Educational Psychology Review* 10, 3 (1998), 251–296. <https://doi.org/10.1023/A:1022193728205>
- [44] C Zhang and D Budgen. 2012. What Do We Know about the Effectiveness of Software Design Patterns? *IEEE Transactions on Software Engineering* 38, 5 (2012), 1213–1231. <https://doi.org/10.1109/TSE.2011.79>