THESIS

CHARACTERISTICS OF INSTRUCTIONAL VIDEOS FOR CONCEPTUAL KNOWLEDGE DEVELOPMENT

P.E. ten Hove
S1360191

University of Twente
Educational Science and Technology
Instructional Technology

Supervisor: Dr. H. van der Meij
Co-supervisor: Dr. J. van der Meij

11-08-2014

UNIVERSITY OF TWENTE.

Abstract

Video is used more and more for instructional goals. To create a high quality design, it is important to know the influencing factors for effective instructional videos. This research investigates the characterization of popular and less popular instructional videos, to discover the differences. The focus is laid on videos with an informative conceptual content, rather than videos that entertain, or support procedural knowledge development. The primary purpose of a conceptual video is to promote deeper understanding of a topic. The study compares various instructional videos [n=75] with conceptual content, presented on the video website YouTube. The videos are classified in three groups (poor, average and good), based on appreciation of viewers (i.e. likes and dislikes) and popularity features (i.e. views and times shared). The groups of videos are compared with a new developed framework based on several influencing design factors, known from the literature. The main distinguishing characteristics of popular videos are the high production quality; the use of a theoretical explanation in combination with illustrative examples; and the use of supportive components (i.e. cues, labels and spoken prompts) to guide the viewer in their learning process. At last, evidence based guidelines are provided for the design of conceptual videos.

Keywords: instructional video, conceptual knowledge development, video design

Acknowledgements

I would like to express my gratitude to my supervisor Dr. Hans van der Meij, for the support in my learning process, the positive comments, and constructive criticism. Furthermore, I would like to thank Dr. Jan van der Meij for the useful comments and remarks on this master thesis. I would like to thank Alinda ten Hove for reviewing my work, being always available, and for her mental support. My thanksgivings also go to Yvonne Luyten - de Thouars and the admission committee EST for the opportunity to attend this master's programme. Finally yet importantly, I want to thank my parents, family, and friends for their support and encouragement throughout my study.

Content

Terms and definitions

Video	Recording of moving visual images made digitally or on videotape
Instruction	'The intentional facilitation of learning toward identified learning goals' (Smith & Ragan, 2005, p. 4)
Instructional video	Video that contains instruction to a large extent
Conceptual video	Instructional videos with as primary purpose to promote deeper understanding of the topic
E-learning	Electronic learning; learning through an electronic interface (e.g. online learning environment with Moodle or blackboard, an online java script course at codeacademy.com)
Blended learning	Combination of traditional education and electronic learning materials
Flipping the classroom	Organization of education: 'knowledge transfer' in the classroom is replaced by videos and any other forms of online instruction. Instruction becomes homework and the homework is done in class.
Animation	A technique that creates the illusion of movement through the projection of a series of still images or frames (at least 12 FPS, frames per second)
Channel	An account on YouTube that uploads videos
Narration	Spoken words (audio)
On-screen text	Printed words (visual)
Graphics	Illustrations (e.g. animation, drawings, charts, graphs, maps, or photos)
HD	High-definition resolution (resolutions of 720 pixels and more)

Introduction

Video exists for 100 years, however in the last decades the use of video for educational reasons has grown exponentially. In particular, the website YouTube with more than 1 billion unique users each month, has contributed to this change. YouTube allows people worldwide to discover, watch and share originally-created videos (YouTube, 2013a). Individuals and companies also use YouTube to share videos to educate other people.

Moran, Seaman, and Tinti-Kane (2011) discovered that YouTube is the most used social media in faculty teaching in U.S. higher education. Eighty percent of the faculty values video use in class, which is the highest percentage of the presented social media. The teachers use videos in class (61%) and for student assignments (32%). Trends in education such as blended learning, flipping the classroom and e-learning make use of those instructional videos. The demand for effective and good quality instructional videos is high.

Hobbyists create many videos shown on the worldwide web; on the other hand, various videos are created for companies or educational institutions. Famous educational video providers are for example Khan Academy and Ted-Ed. Khan Academy is a free online learning platform with thousands of instructional videos for different learning domains. On the other hand, there are also lots and lots of small projects as WePhysics or Biology Professor, that are video channels with course specific videos developed by local teachers. Some videos are viewed millions of times and are very popular, while others are less appreciated. What is it that makes certain instructional videos popular?

This study compares instructional videos to discover characteristics of popular and less popular videos. The focus is on instructional videos that, so we presume, aim for conceptual knowledge development. Conceptual knowledge can be defined as the *'explicit or implicit understanding of the principles that govern a domain and of the interrelations between pieces of knowledge in a domain'* (Rittle-Johnson & Alibali, 1999, p. 175). Conceptual knowledge exists for instance in theories, models, ideas, concepts, definitions and terminologies. In this study, the term conceptual video covers a broad area of videos whose primary purpose is to promote deeper understanding of the topic. The goal is to discover what characterizes popular videos for conceptual knowledge development.

Theoretical framework

Conceptual videos

Instructional videos provide information on a topic with the aim to inform or teach other persons. To narrow the range of videos being considered in this study, the analysis is restricted to videos designed for conceptual knowledge development or simply understanding. To distinguish conceptual videos from other types of instructional videos, it is important to know the differences[1].

In mathematics, the difference between procedural and conceptual information is explained with the difference between skills (*how*) and understanding (*why*) (Baroody, Feil, & Johnson, 2007; Hiebert, 2013; Rittle-Johnson et al., 2001). A related distinction can be made between inform and perform learning goals. Inform goals are lessons that communicate information and perform goals have the aim to build specific skills (Clark & Mayer, 2011). However, these distinctions are general and can be interpreted differently by other people.

Bloom's revised taxonomy of Anderson et al. (2001) distinguishes four types of knowledge: factual, conceptual, procedural and metacognitive knowledge (Krathwohl, 2002). Factual knowledge consists of the basic elements that students must know to understand the subject or to solve problems in the domain. Conceptual knowledge consists of the interrelationships between 'the basic elements within a larger structure that enable them to function together'. Procedural knowledge consists of skills or 'how to do', making use of algorithms, techniques and methods. Metacognitive knowledge consists of 'cognition in general as well as awareness and knowledge of one's own cognition' (Krathwohl, 2002, p. 214).

In this research the conceptual and factual knowledge domain are used to identify conceptual videos. The reason that factual knowledge is used also is because several conceptual videos focus on factual knowledge to explain the subject and to achieve conceptual knowledge. Anderson et al. (2001) specified each knowledge domain with its own knowledge characteristics. In Table 1, the knowledge characteristics of the factual and conceptual knowledge domain are described and clarified with examples.

Table 1 - Structure of the factual and conceptual knowledge dimension based on Bloom's revised taxonomy (Krathwohl, 2002)

Knowledge type	Knowledge of	Examples
Factual	Terminology	Technical vocabulary, symbols and notations, the term 'conjuncture'
	Specific details and elements	Skeleton structure, historical events, elements of a cell
Conceptual	Classifications and categories	Time periods, music styles, forms of child abuse
	Principles and generalizations	Pythagorean theorem, inflation, gravity
	Theories, models, and structures	Maslow's hierarchy of needs, evolution theory, carbon cycle

[1] However, it is important to have in mind that the distinction between conceptual and procedural content is not strict (Rittle-Johnson, Siegler, & Alibali, 2001). In addition, it is known that conceptual learning is more difficult to measure than procedural learning, because understanding is difficult to study from a scientific perspective (Bransford, Brown, & Cocking, 2000).

Analyzing videos

Instructional videos have various features. To analyse instructional videos, the focus can be laid on different aspects. Clearly observable factors are mainly in the physical design, for instance the resolution of the video. Ploetzner and Lowe (2012) provide a structured framework for the analysis of animations (Table 2). Animations are not identical to videos. Nevertheless, there are many similarities. Animations and videos contain visual information, are non-interactive, and are both used to explain a subject. The characteristics of Ploetzner and Lowe (2012) can be used to describe and compare the instructional videos in an analytic manner. However, the cognitive design of the video or animation is not taken into account. The framework is incomplete and needs to be supplemented.

Morain and Swarts (2012) developed a framework for analyzing software tutorials based on Carliner (2000) three-part framework for informational design. The three design levels contain the physical, cognitive and affective design. The physical design is related to the ability to find information, the cognitive (intellectual) design is related to the ability to understand information, and the affective (emotional) design is related to the comfortable feeling in the way the information is presented (Carliner, 2000). Each level of the design contains three objectives (Table 3). The objectives are detailed in an assessment rubric. Three norms are described for each objective, to indicate the video quality as poor, average or good. For example, one of the norms of a good accessible video is that the 'video is cropped to show only task-relevant information'. The framework of Morain and Swarts (2012) is mainly task oriented, because it is developed for tutorials (i.e. 'how to do' videos). The framework needs modification to use it for the analysis of conceptual videos.

Table 2 - Characteristics of animations (Ploetzner & Lowe, 2012)

1	Presentation
1.1	Representations employed
1.1.1	*Visual:* iconic pictures (schematic pictures, realistic pictures, photo realistic pictures), analytic pictures (charts, diagrams, graphs, maps), symbols, formal notations, labels, text
1.1.2	*Auditory:* sound, speech, narration
1.2	*Abstraction:* iconic, abstract
1.3	*Explanatory focus:* behaviour, structure, function
1.4	*Viewer perspective:* single, multiple
1.5	Spatio-temporal arrangement
1.5.1	*Spatial resolution:* constant, variable
1.5.2	Spatial structure
1.5.2.1	Dimensionality: two, three
1.5.2.2	*Organisation:* flat, hierarchical
1.5.3	*Temporal resolution:* discrete, continuous with pauses, continuous with cuts, continuous
1.5.4	Temporal structure
1.5.4.1	*Representation of time:* persistent, implicit, singular
1.5.4.2	*Chronology:* linear, cyclic
1.5.4.3	*Concurrency:* sequential, simultaneous
1.5.4.4	*Organisation:* flat, hierarchical
1.6	*Duration:* presentation time

2.	**User control**
2.1	Time line
2.1.1	Temporal navigation: (re-) start, stop, pause, forward, backward, rewind, fast forward, fast rewind, step by step, go to segment, go to frame
2.1.2	Temporal scaling: change speed
2.2	Presentation
2.2.1	*Appearance:* magnify, change perspective
2.2.2	Information content: zoom, show/hide entities or layers, narration on/off
3	**Scaffolding**
3.1	*Visual:* cues (selection, organisation, integration), written prompts (cognitive, metacognitive)
3.2	*Auditory:* spoken prompts (cognitive, metacognitive)
4	**Configuration**
4.1	*Execution:* single, repeated
4.2	*Setting:* stand-alone, embedded If embedded
4.2.1	*Surroundings:* pictures, animations, video, virtual reality, symbols, formal notations, text, narration, learning tasks, problems, (self-) tests
4.2.2	*Concurrency:* sequential, simultaneous

Table 3 - Rubric for software tutorials (Morain & Swarts, 2012)

Physical Design	
Accessibility	Video allows the viewer to focus on areas of the screen that are relevant to the instruction at hand.
Viewability	Production quality (audio, video, text) is sufficient to make content tolerably watchable.
Timing	Video is paced to make it easy for viewers to follow content.
Cognitive Design	
Accuracy	Content was presented without errors of fact or execution.
Completeness	Content was presented in an organizing superstructure and with sufficient detail so as to be accurately reproduced and broadly applied.
Pertinence	Content was related to the instructional goal, and it had an instructional purpose.
Affective Design	
Confidence	Narrator inspires confidence by presenting self as knowledgeable and skilled. Narrator may also inspire confidence by association with a reputable organization.
Self-Efficacy	Video persuades viewers that they can successfully complete the tasks that are the focus of instruction.
Engagement	Video is designed to interest and motivate users.

Framework

In order to adapt the above-described frameworks to a new model for the analysis of conceptual videos, several test analyses were carried out. Three conceptual videos were elaborated in detail to discover the characteristics of these videos. The results of the test analyses were compared to both frameworks to distinguish which elements could be used or need to be modified. All corresponding elements of the frameworks were implemented in a new structure to analyse instructional videos for conceptual knowledge development. In addition, newly discovered features that were not measured in the frameworks of Morain and Swarts (2012) and Ploetzner and Lowe (2012) were added, such as the use of subtitles. The new structure was used to analyse six new videos, to discover errors and ambiguities. In response to the test phase, the following framework is used for the analysis of conceptual videos (Table 4).

Table 4 - The dimensions in the framework for the analysis of instructional videos

Physical dimension	
Representations	Words and graphics are used to present the instruction
Timing	Video is segmented to make it easy for viewers to follow the content
Production quality	Production quality is sufficient to make content tolerably watchable
Structural dimension	
Structure	Content is presented in an organized structure
Coherency	Visual and audio is related to each other and the instructional goal
Extraneous materials	Unnecessary materials are excluded
Supportive dimension	
Scaffolding	Video supports viewers in their learning process
Personalisation	Video is made personal, by using conversational style
User control	Video can be navigated by the user

The framework is divided in three dimensions: physical, structural and supportive. The dimensions partially relate to the framework for informational design of Carliner (2000). Each dimension has its own characteristics that are divided into several categories. The whole assessment rubric can be found in Appendix A.

Physical dimension

The characteristics in the physical dimension describe the video from an external point of view. Most characteristics can be obtained objectively, without bearing in mind the content. The physical characteristics are subsumed in the categories representations, timing and production quality.

The *representations* are the words and graphics employed in the video. The representations are divided in two parts: visual (e.g. graphics, text) and auditory (e.g. voice-over, sounds) representations. Elements of the model of Ploetzner and Lowe (2012) are adopted and new elements contained from the test analysis are supplemented.

Within the visual representations, a distinction is made between dynamic graphics (i.e. realistic video, animation) and static illustrations. Static illustrations can be iconic pictures (e.g. schematic pictures, realistic pictures, photo realistic pictures) or analytic pictures (e.g. charts,

diagrams, graphs, maps). Visualizations can also include text used as subtitles or to present symbols or formal notations. Subtitles can be implemented in the video itself. On the other hand, YouTube supports the option to enable subtitles. The subtitles can be automatically generated or implemented by the administrator. In addition, subtitles in other languages can be added. If a video only has automatically generated subtitles, it is counted as a video with missing subtitles.

Auditory representations can be characterized by narration (i.e. voice-over that explains the subject matter), speech (i.e. verbalizations that are part of the video), sounds (i.e. non-verbal audio associated with elements in the video), and music. Since videos can consist of different types of representations, the characteristics described are all optional rather than a strict choice.

The *timing* is related to the norm that the video is paced to make it easy for viewers to follow the content (Morain & Swarts, 2012). This can contain the duration of the video or the segmentation. Segmentation is the division of the video in manageable pieces (Clark & Mayer, 2011). Hereby it is important that the narration is natural (i.e. the instruction is not too slow or too quick) and that natural pauses are included. Natural breaks are extended from two to five seconds to allow the viewer to pause and reflect (Morain & Swarts, 2012; van der Meij & van der Meij, 2013).

The *production quality* needs to be sufficient to make the content tolerably watchable (Morain & Swarts, 2012). The production quality of the visualisation can be described with the resolution in pixels (e.g. 720p) and the dimension (i.e. 2D or 3D; Ploetzner & Lowe, 2012). However, in the production process low quality images can be used, while the video is exported in high definition (HD). Low quality visuals are blurred or pixelated. The production quality of the audio is good when the sound is clear. Mechanical noise is a result of poor audio quality. Other noises as breathing or coughing are classified under extraneous audio.

Structural dimension

The characteristics in the structural dimension describe the content of the video and the way information is presented. In comparison with the physical dimension, the subject is taken into account and the characterisation is interpretative. The structural characteristics are subsumed in the categories structure, coherency and extraneous materials.

The content needs to be presented in an organized *structure* (Morain & Swarts, 2012). A typical instruction consists of an introduction, body, conclusion and assessment (Smith & Ragan, 2005). Whether this is the same for instructional videos is questionable. To discover the structure of conceptual instructional videos five questions are asked. 1) Is the content of the video introduced? 2) Is the goal of the instruction given? (Morain & Swarts, 2012) 3) Is the subject theoretically explained? 4) Is an illustrative example given? (van der Meij & van der Meij, 2013) 5) Does the video end with a summary or conclusion? Since videos are non-interactive, questions about assessments are left out.

Coherency means that visual and audio are related to each other and to the instructional goal. Words need to be in line with the visualisation of the video. So, the audio and the graphics correspond to each other and are synchronically presented (Clark & Mayer, 2011). To ensure accessibility the title needs to be drafted carefully (van der Meij & van der Meij, 2013). A carefully drafted title can be discovered if the title explains the content of the video, thus the content corresponds to the title. Finally, the content needs to be relevant for the instructional goal (Morain & Swarts, 2012). Sometimes the goal is not clearly presented. Then the question is if the content is relevant for the instructed subject matter.

Extraneous materials are materials that does not support the instructional goal (Clark & Mayer, 2011). Extraneous elements can occur in the visuals (e.g. irrelevant pictures, mouse

movements), words (e.g. lengthy sentences, non-essential text) and audio (e.g. coughing, loud background music) (Clark & Mayer, 2011).

Supportive dimension

The characteristics of the supportive dimension describe the supportive elements of the video. The supportive characteristics are subsumed in the categories scaffolding, personalisation and user control.

Scaffolding is the cognitive processing support in the learning process of the viewer (Smith & Ragan, 2005). Scaffolds can be visually presented as cues (e.g. arrows, marks) that guide the viewer's visual attention. Scaffolds can also be prompts (e.g. questions, requests) that intend to facilitate the viewer's cognitive process. Prompts can be presented visually (written) and auditory (spoken)(Ploetzner & Lowe, 2012). In addition labels (e.g. titles, key terms) can help for memory support (Clark & Mayer, 2011).

To engage the viewer, *personalisation* of the video is required. Video is made personal, by using conversational style rather than formal style (Clark & Mayer, 2011). Personal aspects are for example by using words like 'you' and 'I' (Clark & Mayer, 2011). Personalisation can also be created with virtual agents or a presenter who is video-recorded. Virtual agents, also called pedagogic agents, are on-screen characters who help to guide the learning process during the video (Clark & Mayer, 2011).

User control gives learners the control over the time line and over the presentation of information in the video (Ploetzner & Lowe, 2012). The videos provided on YouTube can be controlled with the same functions. For that reason, a general description of the user control functions on YouTube is presented in the result section.

Research goal

To answer the question what characterizes popular videos for conceptual knowledge development, several questions need to be answered. To structure the research process, each research part has its own sub questions. Before videos can be analysed, three concepts should be clear: (1) how do you define instructional video for conceptual knowledge development; (2) how do you measure popularity; and (3) which video characteristics need to be measured and how do you measure these characteristics. This last question has mainly been answered by means of the above framework.

After the analysis, the characteristics of instructional videos can be made clear. The characterization of videos contains two ways: (1) what does a general conceptual video look like; and (2) what are the core components of good, average and poor videos, to distinguish differences between popularity. On account of the characteristics of conceptual videos, an application to the practice can be made through guidelines for video designers.

Method

Video sampling

The selection of videos for analysis consisted of three steps: (1) selection of instructional videos for conceptual knowledge development, (2) classification of video popularity, (3) selection of most viewed videos in each category.

Step 1 - Selection of instructional videos for conceptual knowledge development

The website YouTube.com is used to search for instructional videos. Incognito mode is used to avoid personalized search results. Incognito mode, or private browsing, is a feature in web browsers that enable neutral searches that are not influenced by prior browsing history, networks, or friends' recommendations, because information as cookies are not stored or used.

To avoid data biased toward one form of instructional videos for conceptual knowledge development, videos with varied types of learning outcomes were selected. The factual and conceptual knowledge domains of Bloom's revised taxonomy of Anderson et al. (2001) were used to characterize the instructional videos. More information about this model and its characteristics can be found on Page 7 and Table 1. Instructional videos that contain one of the following knowledge characteristics were selected: (1) terminology, (2) specific details and elements, (3) classifications and categories, (4) principles and generalizations, (5) theories, models, and structures. Some instructional videos contain more than one characteristic. The most occurred or most clearly observable characteristic is chosen as main dimension. A condition of the videos used with factual knowledge is that these videos ultimately aim to promote a better understanding of the subject instead of promoting procedural knowledge.

Combinations of general search terms such as ‘‘explanation’’, ‘‘understanding", ‘‘why", and search terms related to the five knowledge types such as "terms", "principle of", "structure", "categories", "types of" etc. were used. The search results on YouTube were filtered with the option 'view count', to prevent finding only popular videos. The large number of hits was screened based on their titles and screenshots. When in doubt of whether to include a video, the video was watched completely or partially.

At first, the video needed to be instructional, but how do you define an instructional video? An instructional video is designed to teach a particular subject. An important characteristic of an instructional video is the explanation. However, not all videos that explain a subject are produced as instructional videos, for example a recording of a lecture or a conference. Lecture-based or ‘substitutional’ videos are recordings of an entire lecture that can be reviewed instead of or after a face-to-face meeting (Kay, 2012). The videos are excluded, because the instruction is not adapted to educate the viewer of the video, but the audience in the hall. Lecture-based videos or conference recordings can be recognized by their public. Another type of videos with an explanation that are excluded are online-published documentaries or television programs. A documentary or television program is sometimes produced to educate the viewer; however, the main reason of television is entertainment.

Secondly, several conditions were checked. A minimum of 1000 views and 25 ratings is maintained and the video needed to be more than one month online, to ensure a reliable picture of the popularity. The only few exceptions for 25 ratings were for videos classified later on as 'poor', which generally have low rating counts. Only videos in the languages English or Dutch were included. The length needed to be between 0 and 30 minutes, because of two reasons. Firstly, in order to prevent an unnecessarily long analysis process. Secondly, most of the longer videos on YouTube are recorded lectures or documentaries. Videos clearly produced for

children (12 year and younger) are excluded, because the target group is adults and young adults (age 12 year and older). For example, when words as 'for kids' and 'children' occurred in the title, the video was excluded.

To define which video promotes conceptual knowledge development, the characteristics of the knowledge types of Bloom's revised taxonomy are used. When the video met all the above-described conditions, the knowledge type of the learning outcome was determined, based on the title and the first impression of the video. Of each category of knowledge types, around the same amount of videos were selected, until a total database of 250 videos resulted. In the total database, no more than five videos per channel and five per subject were recorded, to ensure a broad range of instructional videos. A channel is the account that uploaded the video and a subject is the topic of the video such as 'gravity'.

The selected videos were downloaded with the program 'Free YouTube Download v. 3.2.29. build 303' and saved as .mp4 with each a unique ID. Each fixed information (e.g. the channel, the upload date and the duration) and time-varying information (e.g. view count, ratings and times shared) of the videos were recorded and tabulated using Microsoft Excel. An overview of the external characteristics is shown in Appendix B, and the specification of the derivation of data is described in the codebook (Appendix E).

Step 2- Classification of video popularity

The selected videos in the total database were categorized based on a quality score in three categories: poor, average and good. In the research of Morain and Swarts (2012) about software tutorials, videos rated 3.5-5.0 stars, rated 2.6-3.4 stars, and rated 0-2.5 stars were selected to be good, average and poor. However, this rating approach could not be adopted. The current videos on YouTube are not valued with stars (1-5) anymore, but with thumps up and down. A thump up relates to 'like' and thumps down to 'dislike' (Figure 1). This type of rating is less specific and less reliable than the rating with stars. The new indicator is based on the appreciation of viewers in combination with popularity characteristics, in order to gain a clearer indication of the video.

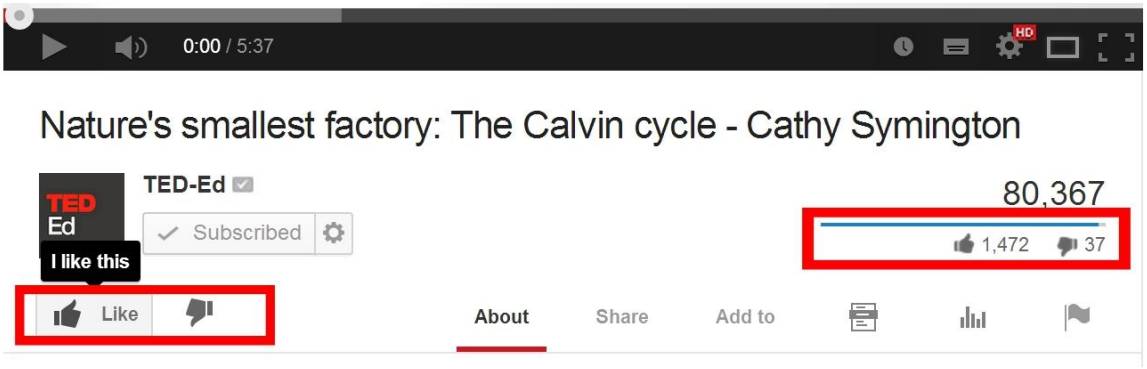


Figure 1 - Rating through 'Likes' and 'Dislikes'

A formula is created to indicate new quality scores to categorize the videos as poor, average and good. The quality score is calculated based on the amount of likes (L), dislikes (D), views (V), and times shared (S) of the video. An overview of the data types and formula is shown in Table 5. Since the distribution of data is quite broad (e.g. ten to millions of viewers), percentages (e.g. likes in relation to views) deliver such a small number, which is not in proportion. That is why the data is categorized in different categories (1-5).

At first, the amount of likes is compared with the total number of likes and dislikes (R). In the formula, the dislikes are counted twice. One of the reasons is that the differences in ratio were very small. Another reason is that the probability that someone dislikes a video would be low. For instance, the change that the viewer skips the video before providing his opinion is much higher. That is why the dislikes are counted heavier than likes. Secondly, the amount of views is categorized (CV). Hereby the videos with less than 1000 views are removed. Thirdly, the number of times the video is shared is categorized in five groups (CS). The three scores are summed up and divided by 4 (2R+CV+CS)/4. The ratio of likes and dislikes is counted twice in the final formula. The reason for that is that the amount of likes and dislikes tells us more about the quality than for instance the amount of views.

The outcome of the total ratio (TR, or called quality score) is between the number 1 and 5. To distinguish the different quality groups the total ratio is divided by three to provide three groups. Videos rated 3.7-5.0, rated 2.4-3.6, and rated 0-2.3 are selected to be good, average and poor.

Table 5 - Used data and formula to categorize videos as poor, average and good

Available data	New data	Formula	Outcome	Grouping	Category
V = Views L = Likes D = Dislikes S = Times Shared T = Length of video M = mean time watched[2]	R = Ratio Likes/Dislikes*2	L/(L+2D)*100	0-100%	0-60%	1
				60-70%	2
				70-80%	3
	CR = Category Ratio	-		80-90%	4
				90-100%	5
	CV = Category Views	-	≥ 0	0-1000 (Not used)	-
				1.000-10.000	1
				10.000-100.000	2
				100.000-1.000.000	3
				1.000.000-10.000.000	4
				>10.000.000	5
	CS = Category Shared	-	≥ 0	0-1	1
				1-10	2
				10-100	3
				100-1.000	4
				> 1.000	5
TR = Total ratio		(2R+CV+CS)/4	1-5	**1-2.3**	**Poor**
				2.4-3.6	**Average**
				3.7-5	**Good**

Step 3 - Selection of most viewed videos in each category.

A cluster sampling is used to select instructional videos for the analysis. All 250 videos in the database were ranked in order of number of views. In each category of knowledge dimension and quality group, the first 5 videos were selected. With 5 knowledge types and 3 quality groups, a total of 75 videos were used for further analysis.

[2] YouTube provides since 2013 more features to analyze your videos. Older videos do not contain of this information, such as the statistic about the mean time watched. For that reason, the statistic is not used for the classification of videos.

Instruments

To analyse the selected videos, several characteristics are measured. Two instruments are used for the characterisation of the instructional videos. The first instrument consists of the external properties, which is presented in Appendix B. The external properties consist of 14 items that are obtained from fixed data presented on YouTube, below the video. The second instrument consists of the new framework presented in the previous section. The framework consists of 36 items divided into the three dimensions, based on features of instructional videos known from the literature. An overview is presented in Appendix A.

A codebook is created to clarify the instruments with each item and its measurement (Appendix E). Each item in the codebook is described in detail with indicators and examples.

Data analysis

Framework

To analyse the data obtained from the videos and the framework, statistical analysis using SPSS is employed. In order to test which characteristics distinguish poor, average and good conceptual videos, chi-square tests evaluate the differences between the videos in the three popularity groups (poor, average and good) and the different variables. Besides the existing variables, some variables are combined or recoded into new variables. Most variables are nominal, with few exceptions for ordinal data (e.g. narration speed) or scale (e.g. video length). Chi-square (χ^2) tests revealed the significant differences between the ordinal and nominal variables in comparison with the popularity of the video. In some cases the data did not meet the conditions of the chi-square test. In those cases, groups were merged, such as the answer 'yes' and 'sometimes' and the popularity groups 'average' and 'poor'. For effect size, Pearson's chi-square statistic or Fisher's exact test is reported. All analyses are two sided with alpha set at 0.05.

For describing the results, the items coded 2 (yes) and 1 (sometimes or unclear) are merged in the physical and supportive dimension. This because answering 'sometimes' means that the characteristic is represented in the video although occasionally. For example, a video contains a small part of animation. In the structural dimension, the coding 'sometimes' was more often used when the determination of the characteristic was unclear. For example, it was unclear if the video contains an introduction, because the story continued without clear transitions. That is why the results of the structural dimension represent only the times answering 'yes'.

External properties

Besides the measured variables in the framework, other external variables were reported (Appendix B). To compare the scale variables with popularity, an analysis of variance (ANOVA) is used. All external variables show a relation with popularity, $p < 0.05$. Three relationships between the external properties and popularity are important to mention. First, a statistically significant effect between the days online and the original quality score was measured, F (20, 54) = 4.116, $p < .001$. With a mean age of 2 years, good videos were relatively newer, in comparison with average and poor videos. Average and poor videos were on average more than 3 and half years online. Second, as can be expected a relation between popularity and channel members is measured, F (2, 72) = 24.134, $p < 0.001$. The same applies to the percentage of the video watched and popularity, F (2, 17) = 8.567, $p = 0.003$.

Reliability

Cohen's (weighted) Kappa (κ) is used to prove the inter-rater reliability of the framework. According to the commonly used interpretation-scale of kappa, $\kappa > 0.61$ is used as indicator for a substantial agreement. The researcher and a second academician applied the codebook to characterize six randomly selected videos. The results of both coders were compared between all videos. The results showed sufficient agreement between the two coders. The overall mean score was $\kappa = .661$ (95% CI, .567 to .756), $p < .001$.

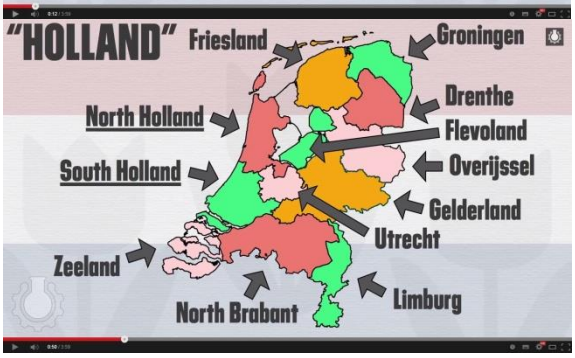
In addition, the results of each item were separately compared between the two coders, to discover which items needed extra attention. The items iconic pictures (C), narration speed (N), noisy audio (S), illustrative example (W) and extraneous audio (AE) showed an extremely low agreement ($\kappa < 0.20$). The items formal notations (F), theoretical explanation (V), natural pauses (O) and style (AI) showed a fair agreement (κ =0.20-0.40). The remaining 27 items showed a moderate to very good agreement (κ =0.40-1.00). The list of all results is presented in Appendix C.

Results & Discussion

The dataset consists of 75 videos extracted from 62 unique YouTube channels and with 72 unique topics. The dataset includes 70 English videos (93,3%) and 5 Dutch (6,7%). The characterization of the videos will be done by describing a general conceptual video first. Thereafter the distinguishing characteristics of the three popularity groups will be elaborated. An overview of the main results is presented in Appendix C.

Instructional videos for conceptual knowledge development

While it is difficult to find an instructional video that exemplifies all of the typical characteristics of instructional videos for conceptual knowledge development, one titled "Holland vs. the Netherlands" comes close on most accounts. It is certainly good enough to point to key communication design features to outline the type of videos discussed. The video uploaded on December 21, 2012 explains the difference between the terms 'Holland' and 'the Netherlands' by delving into the structures of the country. The video is categorized as good. The video contains various features of conceptual videos. Several characteristics of the knowledge types of Bloom's revised taxonomy can be recognized: explanation of terminology, specific details and elements, categories and structures. In Figure 2, some fragments of the video are presented to illustrate the results.

2.1	*Narration* Welcome to the Great nation of Holland: where the tulips grow, the windmills turn, the breakfast is chocolaty, the people industrious, and the sea tries to drown it all. Except, this country isn't Holland. It's time for: The Difference Between Holland, the Netherlands (and a whole lot more) *Timing* 0.00-0.12 (12 seconds)	
2.2	*Narration* * Noord (North) Holland and * Zuid (South) Holland. These provinces make calling the Netherlands 'Holland' like calling the United States 'Dakota'. Though unlike the Dakotas, which are mostly empty, save for the occasional Jackalope, the two Holland's are the most populated provinces and have some of the biggest attractions like Amsterdam and Keukenhof. *Timing* 0.46-1.02 (16 seconds)	
2.3	*Narration* But why does the Kingdom of the Netherlands reach to the Caribbean anyway? Because, Empire. In the 1600s the Dutch, always looking to expand business, laid their hands on every valuable port they could. *Timing* 2.29-2.39 (10 seconds)	

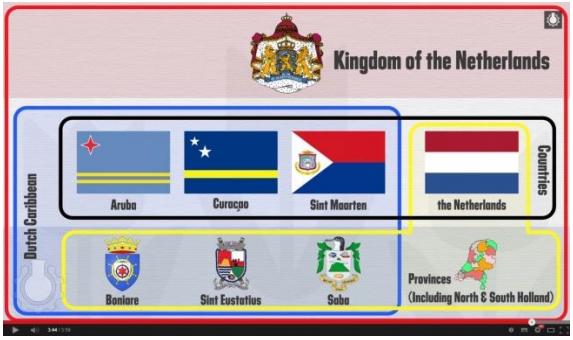
2.4	*Narration* So in the end, there are 6 Caribbean islands, four countries, twelve provinces, two Holland's, two Netherlands **and one kingdom, all Dutch.** *Timing* 3.37-3.45 (8 seconds)	

Figure 2 –Fragments of the conceptual instructional video 'Holland vs. the Netherlands' created by CGPGrey (2012)

Physical dimension

Representations

In general, conceptual videos make use of different types of visual materials. Half of the videos (52%) use a combination of materials, others use one specific way to present information (Table 6). No clear preference is seen between the graphics. More variation is seen in the use of textual representations. Symbols and formal notations are not used a lot (21 %) and less than half of the videos makes use of optional or pertinent subtitles (44,6%). In the use of audio materials, there is preference for narration (62,7%), which is in line with the use of a video-recorded presenter. In total, the videos mostly use audio to present words (89,3%). Few videos use sounds, while music is regularly used in the introduction of the video (24%) or as background music throughout the video (40%). In our example video, the used representations consist of a combination of various pictures (e.g. Figure 2), optional subtitles, narration and discrete background music.

Timing

The average length of the videos is 3 minutes and 50 seconds, which is almost the same as the total length of the example video with 3 minutes and 59 seconds. In general, most videos are spoken at natural speed (69,1%) with several natural pauses longer than two seconds (76,5%). This is in contrast with the example video, were the narration speed of the video is quite fast, and natural pauses are less used.

Production Quality

In general, no clear preference is seen in the use of resolution, although HD is used the most (49,4%). The resolution of the example video is also uploaded in HD (1080p). Two-dimensional materials are used mostly in instructional videos (80%), although sometimes in combination with three-dimensional materials (35,0%). In the example video, two-dimensional materials are used, which are represented in flat images that do not represent depth (Figure 2). Half of the videos show no blurred visuals (53,4%), and the videos are often free of mechanical noise in the audio (75,4%). The example video is consistent with the overall results. The visual materials are clear and bright, and the audio is noise-free.

Structural dimension

Structure

When exploring the structure of the video, it can be seen that regularly an introduction is given (57,3%), such as in this example by presenting the topic by a small story and a title page (Figure 2.1). However, in most videos no clear goal is presented (75,7%), despite the fact that it is important to start with a goal or a set of objectives (Swarts, 2012).

The core of the video consists of a theoretical explanation (78,4%) or an illustrative example (68,9%). Half of the videos uses a combination of theory and practice (51,4%) as in our example. In the example, historical events are used for the theoretically substantiation of the name 'Kingdom of the Netherlands' used in other countries (Figure 2.3). In addition, examples related to the viewer's background are used to explain the subject. For example, the comparison between the provinces of North and South Holland and the American states North and South Dakota (Figure 2.2), which is identifiable for the American viewer.

The closure of the video consists in some cases of a summary or conclusion (35,1%). The example video concludes with a summary of the major differences in names of the Netherlands, by providing an overview (Figure 2.4). For most of the videos, it was difficult to recognize the structure of the video. This is in contrast to what is recommended by Swarts (2012), who recommends making visible the sections of the video.

Coherency

Instructional videos show coherence between the materials. Video and audio is related to each other (91,7%) and presented synchronically (88,9%). For instance, at the moment the narrator is talking about trading in 1600, a picture is shown of traders in that time (Figure 2.3). In most of the videos, represents the title the content of the video (72%). In addition, the content of the video contributes to the explanation of the subject or goal (88%). The title of the example video is 'Holland vs. the Netherlands', which represents the content of the video. All information given in the video contributes to the explanation of the differences between those names. For example, in Figure 2.2 an overview of all provinces of the Netherlands is presented. This overview helps the viewer to understand that 'Holland' consists of two provinces of the country the Netherlands. The Netherlands, however, is actually much bigger than just the country and expires even to the Caribbean (Figure 2.3).

Extraneous materials

In general, relatively few errors are seen in conceptual videos. Extraneous materials such as extraneous visuals (22,7%), words (13,5%) and audio (17,8%) are scarcely seen. The example video is presented without extraneous materials. For example, words as 'ehm' and noise such as mouse clicks are not discovered.

Supportive dimension

Scaffolding

Several components are used in conceptual videos to support the viewers in their learning process. Cues are used to guide the viewer's visual attention (52%). For example, arrows are used to pinpoint the provinces (Figure 2.2) and coloured frames are used to clarify structures (Figure 2.4). What is even more frequently used, are labels (73,4%). Labels present for instance names of the provinces (Figure 2.2) or countries (Figure 2.4). Another way of supporting the viewer in their learning process is by the use of prompts. Prompts are questions or requests that can be presented written and spoken (Ploetzner & Lowe, 2012). Written prompts are not

frequently used (25,3%), while several videos consist of spoken prompts (54,6%). Spoken prompts are primarily seen in questions that trigger the viewer to reflect. For instance, *'But why does the Kingdom of the Netherlands reach to the Caribbean anyway?'* (Figure 2.3).

Personalisation

Most of the videos show no clear preference for narration style (60,8%). For instance, the style of the example video is not clearly conversational or formal. Words as 'you' and 'I' are not used, however formal language and complex sentences are avoided. Personalisation through visuals is rarely used in instructional videos. Most videos do not use a video-recorded presenter (26,7%) or an animated agent (2,7%). The same can be seen in the example that not uses an on-screen presenter.

User control on YouTube

Enabling user control is an invitation for the user to become an active learner, by influencing the playing of the video (van der Meij & van der Meij, 2013). YouTube offers quite a lot of user control options. Several options help the viewer to navigate through the video. A play/pause button is presented, which changes into a rewind button after ending the video. In addition, the user can pause and play the video with the spacebar. To jump to another point in the video, the user can move the button on the playback control bar. While moving the button or by just moving the mouse above the playback control bar, the time and a preview shot of the video is shown. The screen can be magnified with two buttons. The first button changes the settings between the 'default view' (i.e. the video is displayed small with a list of recommended videos at the right side) and the 'cinema mode' (i.e. widescreen). The second button changes the video between half screen and full screen. The volume can be adjusted using a slider or muted by pressing the same button. With the settings button, the resolution of the video can be changed into a lower quality and the speed of the video can be slowed down or accelerated (0.25 to 2 times faster). Furthermore, annotations and captions can be switched off and on. A more detailed description with examples is provided in the codebook (Appendix E). The functions described above are available for browsing with Google Chrome, according to the website www.youtube.com at July 8, 2014. Not all functions are available in all browsers and devices.

Video characteristics of the three popularity groups

To discover what characteristics vary between popularity, the results are compared with the groups good, average, and poor. Several characteristics of instructional videos differ between the three conditions. Quite a lot of the measured variables show a pattern in the groups of popularity, however not all patterns are statistically significant. Table 7 shows an overview of all characteristics that differ significantly between the popularity groups.

Table 7 - Percentages of the significant characteristics of conceptual videos

Item			Percentage (%)			*Chi-square*
ID	*Variable name*		*Good* (*N* = 25)	*Average* (*N* = 25)	*Poor N* = 25)	*p-value*
Physical Dimension						
C	Iconic pictures		**76,0**	36,0	44,0	.025
D	Analytic pictures		**68,0**	28,0	16,0	.002
A-D	Combination of graphics		**76,0**	40,0	40,0	.013
H-J	On-screen text instead of narration		4,0	4,0	**24,0**	.019
I-L	Audio	*Narration or speech*	8,0	**52,0**	28,0	.000
		Music or sounds	4,0	4,0	**28,0**	
		Both	**88,0**	44,0	44,0	
H	Subtitles	*No subtitles*	20,0	**80,0**	66,7	.000
		Optional subtitles	**76,0**	16,0	4,2	
		Pertinent subtitles	4,0	4,0	**29,2**	
N	Narration	*Slow*	0,0	8,3	**15,0**	
	speed	*Natural*	50,0	**87,5**	80,0	.000
		Fast	**50,0**	4,2	5,0	
O	Natural pauses are included		54,2	**95,8**	80,0	.004
P	Resolution	*240p*	4,0	16,0	**20,0**	.000
		360p	8,0	12,0	**36,0**	
		480p	4,0	**40,0**	12,0	
		720p (HD)	20,0	**24,0**	**24,0**	
		1080p (HD)	**64,0**	8,0	8,0	
Q	Dimensionality	*2D*	**56,0**	44,0	**56,0**	.031
		Both 2D and 3D	4,0	8,0	**28,0**	
		3D	40,0	**48,0**	16,0	
R	Visual Blurred		12,0	56,0	**72,0**	.000
S	Noisy Audio		4,2	**36,0**	33,3	.003*
Structural Dimension						
W	Illustrative example		**84,0**	64,0	58,3	.038*
T-W	Combination Theory/Practice		**72,0**	48,0	33,3	.024
AE	Extraneous Audio		4,2	24,0	**25,0**	.027*
Supportive Dimension						
AF	Cues		**68,0**	52,0	36,0	.042*
G	Labels		**92,0**	76,0	52,0	.003
AH	Spoken prompts		**70,8**	45,8	44,5	.024

** Fisher's exact test: average and poor videos combined*

Physical dimension

Representations

The use of *graphics* differ by popularity, χ^2 (7, N = 68) = 18.718, p = .009. No difference is found between realistic video and animation (p > .05), while iconic and analytic pictures show variation by popularity (p < .05). Table 6 presents an overview of the graphics in the dataset.

Table 6 - Graphics presented in the dataset in percentages (%) of total

Graphics	Total % (*N* = 75)	Good (*N* = 25)	Average (*N* = 25)	Poor (*N* = 25)
Dynamic graphics	77,3	80,0	76,0	76,0
Realistic video	41,3	48,0	44,0	32,0
Animation	52,0	56,0	44,0	56,0
Static illustrations	64,0	92,0	52,0	48,0
Iconic pictures	52,0	76,0	36,0	44,0
Analytic pictures	37,3	68,0	28,0	16,0
No graphics	2,7	0,0	4,0	4,0
Combinations of materials	52,0	76,0	40,0	40,0

Good videos often use more iconic pictures in comparison to average and poor videos (72% versus 36% and 44%, $p = 0.003$, Fisher's exact test), the same applies to analytic pictures (60% versus 28% and 16%, $p = 0.001$, Fisher's exact test). Another result is that the use of combinations of graphics differ by popularity, χ^2 (2, $N = 73$) = 8.654, $p = .013$. Good videos often use more combinations of visual materials in comparison with average and poor videos (76% versus 40%, $p = 0.003$, Fisher's exact test). In the calculation the use of analytic and iconic pictures were combined to represent the use of pictures.

The graphics in good videos are used to illustrate as much as possible of what they are telling, no matter the type of graphic. Almost all good videos use pictures. From the literature is known that simpler graphics lead to better learning than a more realistic or complex visual (Clark & Mayer, 2011). From instructional design it is known that images are a powerful for processing information, especially for learning facts and lists (Smith & Ragan, 2005). In addition, pictures are easy to use and are less time-consuming to produce than dynamic graphics. Therefore, this factor is not depending on the professionalism of the video producer and the available money.

In terms of *textual representations*, a relationship with popularity is found with the variable 'subtitles' (χ^2 (4, $N = 71$) = 40.005, $p < .001$). Good videos are presented more often with the possibility to add different subtitles (76% versus 16% and 4,2%, $p < 0.001$, Fisher's exact test). Poor videos are presented sometimes with pertinent subtitles or present text instead of audio (29,2 versus 4%, $p < 0.001$, Fisher's exact test). The use of text presented in the narration or on-screen differs by popularity χ^2 (2, $N = 73$) = 7.876, $p = .019$. Good and average videos present words less as on-screen text than poor videos (4% versus 24%, $p = 0.038$, Fisher's exact test).

That subtitles are used only as optional in good videos, can be explained by the *redundancy principle* of Clark and Mayer (2011). It is important to reduce on-screen text, because this needs to be processed and takes cognitive load of the visual channel. Subtitles could only be used in specific situations, when no graphics are presented or when the learner is no native speaker (Clark & Mayer, 2011). Therefore, optional subtitles are only relevant for the design of videos for the international market. In other situations it is important to use audio narration to present words (van der Meij & van der Meij, 2013), which is summarized in the *modality principle* (Clark & Mayer, 2011). Results of this study confirm this principle, since good videos usually use audio narration in contrast to poor videos.

The analysed videos do not differ significantly in the use of the different specified *audio materials* ($p > .05$). Although, a difference is found between the use of only narration or speech, music or sounds, or a combination of both, (χ^2 (2, $N = 73$) = 21.773, $p < .001$). Good videos more

often used a combination of these audio materials (88%) than average and poor videos (44%, $p < 0.001$, Fisher's exact test). The audio of poor videos sometimes only consists of music or sounds (28% versus 4%, $p = 0.024$, Fisher's exact test). In most of these cases, words were presented as text, in few cases no words were used at all for the explanation.

Poor videos contain frequently music and sounds. However, good videos contain also regularly music and sounds, while it is known that music and sounds may overload the working memory and can hinder learning (Clark & Mayer, 2011). Some studies find contradictions in the supportive role of music in virtual educational environments (Fassbender, Richards, Bilgin, Thompson, & Heiden, 2012). More research is needed to determine if there are situations that sounds have more advantages than disadvantages (Clark & Mayer, 2011). For example, the use of music in an example of a good video in the next section seems to contribute to the explanation and attraction of the video (Figure 7). According to the *scissors of Wember*, image and sound must be consistent with each other (Wember, 1976). When music and visuals are in line, then the music is the bearer of the storyline (Bueters, 2002). In addition, it is important to take into account that music has copyrights. This can influence the use of music in instructional videos, because YouTube blocks your video when you use music with copyright.

Timing

The timing used in the videos show a relationship with popularity (Figure 3). Popular videos seem to be slightly longer and display more variation in length (good: $M = 0{:}06{:}55$, $SD = 0{:}04{:}38$; average: $M = 0{:}05{:}07$, $SD = 0{:}03{:}30$; poor: $M = 0{:}04{:}42$, $SD = 0{:}03{:}27$). The difference is statistically significant between good and average/poor videos combined, $F(1, 73) = 4.462$, $p = 0.038$. The speed of the narration also differs between the different popularity groups, $\chi^2(3, N = 72) = 26.219$, $p < .001$. A Spearman's rank-order correlation was run to determine the relationship between the narration speed and the original quality score. There was a moderate, positive correlation between narration speed and popularity, which was statistically significant, $r_s(73) = .415$, $p < .001$. In addition, the use of natural pauses differ by popularity, $\chi^2(2, N = 73) = 18.984$, $p = .004$. Natural pauses were less used in good videos (54,2%) in comparison with average (95,8%) and poor videos (80%, $p < 0.001$, Fisher's exact test).

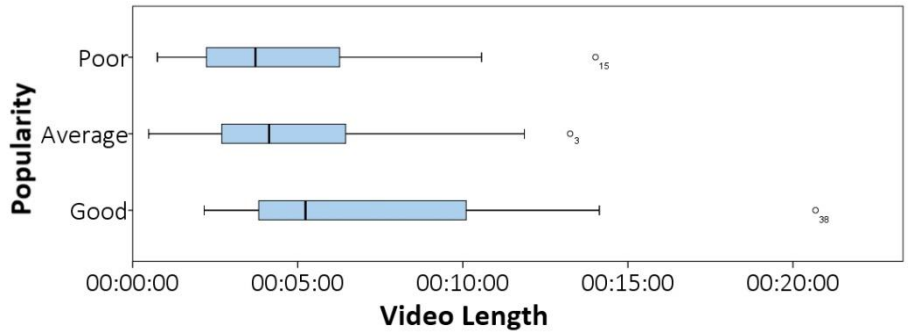


Figure 3 - Video length versus popularity

These findings are quite remarkable, because the literature told us other things so far. Literature about procedural videos recommends that videos need to be as short as possible (van der Meij & van der Meij, 2013), because shorter videos are more engaging than longer videos (Ruedlinger, 2012). Different advices of length are made, for example the video needs to be shorter than 6 minutes (Guo, Kim, & Rubin, 2014). More than half of the good videos in this data set are longer than that, and yet they are more popular than the shorter videos in the groups

average and poor. One reason may be that the introduction of good videos is more attractive, so that the viewer keeps watching. This because dropout rates change quite dramatically at the beginning of a video (Kim et al., 2014). After the intro of the video, there is not a major variation in engagement for a 4-minute versus a 10-minute video (Ruedlinger, 2012). In the study of Guo et al. (2014) the average percentage of the video being viewed extremely drops out with videos longer than 9 minutes. Apparently, the longer videos in the category good are such interesting videos, that viewers keep watching. Another possible reason is that conceptual videos need more time to promote the deeper understanding of the topic. However, this needs to be explored in further research.

Another point that does not match the literature is that narration speed of good videos was experienced much faster and natural pauses were less used in comparison with average and poor videos. From the literature it is known that moments of reflection increase learning from instructional video (van der Meij & van der Meij, 2013). The advice is to speak instructions not too quickly and to include natural breaks to allow the viewer to pause (Morain & Swarts, 2012). One reason that poor videos use more natural pauses is that they are less well prepared and consist of mistakes, which causes more pauses in the video. In addition, the speed of the narration in good videos is faster and allows less space for natural pauses longer than the measured two seconds. On the other hand, videos where instructors speak fairly fast and with high enthusiasm are more engaging (Guo et al., 2014). Since especially the introductions were spoken fast, this could contribute to the fact that the introductions of good videos are more attractive. However, it is important to take into account that this study describes what is actually popular and not what is the best for learning. With the rise of video clips, there is a strong tendency to particularly impressive and quick settings with lots of visual effects, which the viewer becomes more and more used to (Bueters, 2002). However, the overload of visual effects is not always in line with the goal of the video (Bueters, 2002). A consequence of fast-talking could be that the viewer gets sucked into the video and keeps watching. The question remains whether this is good for the learning process of the viewer.

Production Quality

The production quality seems to have an influence on popularity or the other way around. First, the resolution of the uploaded videos differs by popularity, χ^2 (4, N = 71) = 38.032, p < .001. A Spearman's rank-order correlation was run to determine the relationship between resolution and the original quality score. There was a moderate, positive correlation between resolution and popularity, which was statistically significant, $r_s(73) = .567$, $p < .001$. In addition, a relation between resolution and the days online was found, F (4, 71) = 18.349, p < .001. Second, the use of two- or three-dimensional materials differ by popularity, χ^2 (4, N = 71) = 10.662, p = .031. No difference was found between good and average videos (p > .05), while poor videos made less use of three-dimensional materials (40% versus 16%, p = 0.029, Fisher's exact test). Third, the display of blurred visuals differs by popularity, χ^2 (2, N = 73) = 24.986, p < .001. Poor (72%) and average (56%) videos more often displayed blurred visuals in comparison with good videos (12%, p = 0.011, Fisher's exact test). Blurred visuals differ also by the resolution of the video, χ^2 (12, N = 63) = 76.441, p < .001. Fourth, temporary to continuous mechanical noise in the audio differs by popularity, χ^2 (2, N = 73) = 8.128, p = .017. Poor (33,3%) and average (36%) videos more often displayed noisy audio in comparison with good videos (4,2%, p = 0.003, Fisher's exact test).

According to the results, production quality is important for the popularity of the video. For example, a high resolution, three-dimensional materials, sharp visuals and clear audio

contribute to a high production quality. However, it would not be logical that the quality influences the popularity, but the other way around. When a video is produced carefully and with lots of effort, the designer will publish the video in good quality. On the other hand, text must be legible to perceive the information from it. In addition, blurry images are often seen as distracting and annoying. Distracting materials guiding the viewer's limited attention away from the relevant material and towards the irrelevant material (Clark & Mayer, 2011; Wember, 1976). Even small mistakes are problematic, because viewers notice these errors (Swarts, 2012). That is why high quality audio and HD quality video are necessary (Swarts, 2012).

The production quality is a precondition for success, because the viewer needs to be able to perceive the information. For example, degraded audio is an obstacle to hearing instructional messages (Swarts, 2012). Nevertheless, a high production quality is no guarantee for success. Some of the poor videos were produced of high quality materials; however, they were rated very poor. After having dug into the comments of the video, it became clear that the content contains obstacles for people. High quality produced videos with low ratings had to deal regularly with sensitive issues or the video provided incorrect information. For example, a video about photosynthesis used citations of the Koran. This interruption did not contribute to the explanation of the topic and is subordinated to personal opinions.

The days online, the resolution and the popularity of videos correlate with each other. Over the years, the quality of recording devices and upload capability has increased. The question is whether the popularity increases by time is caused by the increase of resolution quality or that the instruction quality of videos has increased by time. The latter could be explained by the growth of YouTube and its use by professional companies (YouTube, 2013a).

Structural dimension

Structure

Providing an example or using a story in the explanation of a topic is used more often in good videos in comparison with average and poor (96% versus 84% and 82,3%, $p = 0.038$, Fisher's exact test). Providing an illustrative example in combination with a theoretical explanation shows a larger significant difference between the popularity groups, χ^2 (2, $N = 73$) = 7.498, $p = .024$. Good videos use more often a combination of theory and practice (72%) than average (48%) and poor videos (33,3%). The other structural variables show no statistical relation with popularity ($p > .05$).

Although it is important to keep in mind that the items ‘theoretical explanation’ and ‘illustrative examples’ show a low inter-rater reliability, these results are important. Guidelines on procedural videos advise to strengthen the instruction with practice (van der Meij & van der Meij, 2013). The purpose of procedural videos is to explain procedures and skills, which are closely related to practice. The explanation of theories and concepts in conceptual videos is more theoretical and less directly related to the practice. Nevertheless, the dataset shows different ways to couple theory to the practice in conceptual videos. For example with stories (Figure 4.1), examples related to real life (Figure 2.2 & 4.2) or examples to express something symbolically (Figure 10.2). Experiments or observations (Figure 4.3) were less usual in popular videos and are depending on the topic (e.g. explaining physics). From instructional design theory it is known that providing examples is critical for the information processing (Smith & Ragan, 2005).

Illustrative examples are used in two ways. Firstly, by providing an example and explaining it with theory (Figure 9.4). Secondly, by explaining the theory and giving an example (Figure 4.2). These two strategies are known as the inquiry and expository approach for conceptual learning

in instructional design theories. The inquiry approach starts with examples, which stimulates the learner to discover the underlying conceptual information (Smith & Ragan, 2005). The expository approach starts with the conceptual information, which stimulates the learner to develop his own examples (Smith & Ragan, 2005).

In the research of Magner, Glogger, and Renkl (2014) three types of illustrations that trigger situational interest are distinguished: concreteness, personal relevance and ease of comprehension. Especially illustrative examples that contain concreteness are seen in conceptual videos.

4.1	**Why do competitors do open their stores next to one another?** (De Haan, 2012) *Narration* Imagine that you sell ice cream at the beach... *Timing* 0.43-0.45 (2 seconds)	
4.2	**Biological Molecules** (CrashCourse, 2012a) *Narration* Some foods, especially ones that contain animal protein, have all of the essential amino acids **including this egg** *Timing* 13.20-13.27 (7 seconds)	
4.3	**Water** (CrashCourse, 2012b) *Narration* We're going to do some pretty amazing science right now. You guys are not going to believe this. Ok, you ready? **It floats!** Yeah, I know you're not surprised by this, but you should be, because everything else, when it's solid, is much more dense than when it's liquid, just like gases are much less dense than liquids are. *Timing* 7.46-8.06 (20 seconds)	

Figure 4 - Examples of coupling theory to practice

Notable is, that videos in the category 'good' seem to use more than one theoretical explanation. Diverse sources are used and different aspects of a problem are discussed (e.g. example of a good video at the end of this paragraph, Figure 9). In some average videos the explanation is also theoretically grounded, however this is not related to practice or strengthened with illustrative examples. For example, the video presented in Figure 5 explains what enzymes are and how they work. Throughout the video, several models are used to explain how enzymes work. However, the relation with real life is missing. Providing illustrative examples will improve the explanation. For example, the relation with nutrition and the functioning of our body can be given, to stimulate insight in what enzymes are and what they do.

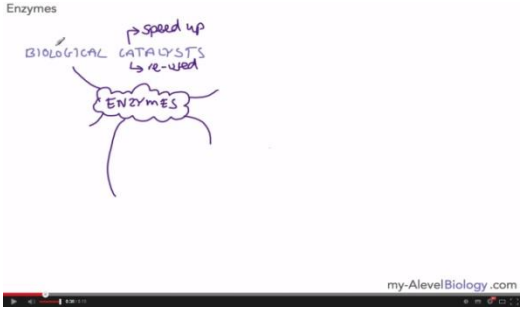
5	**Enzyme Action** (myAlevelBiology, 2013) *Narration* Lets have a quick overview of what enzymes are all about. The first thing is that they are all biological catalysts and by catalysts we mean substances that can speed up chemical reactions. And the other thing about catalysts is that they do not only speed up chemical reactions, but they don't actually get involved in the reaction, that means they can't be re-used. *Timing* 0.12-0.36 (24 seconds)	

Figure 5 - Fragment of an average video explaining enzymes

In this research, no significant differences between the other structural characteristics and popularity were found. A reason for this could be that structural elements such as a summary or conclusion in general are infrequently used in conceptual videos. Although in this research 35,1% of all the videos contain a summary or conclusion, which is not negligible. According to personal contact, it became clear that Ploetzner and Lowe (2012) did not find evidence for the use of a summary or conclusions in animations.

Coherency and Extraneous materials

The coherency variables show no statistical relation with popularity ($p > .05$). The same applies for extraneous materials, except that extraneous audio was more common in poor/average videos combined in comparison with good videos (24,5% versus 4,2%, $p = 0.027$, Fisher's exact test). Extraneous audio could be caused by for example loud music or background noises, such as coughing. However, the low reliability of this item shows that this is experienced differently per person as hindering or not. Even though extraneous sounds may overload the working memory and could distract the viewer from relevant materials (Clark & Mayer, 2011) and thus could be better reduced.

Supportive dimension

Scaffolding

Three forms of support are important in conceptual videos; cues, labels and spoken prompts. Cues were used more in good videos in comparison with average and poor videos (68% versus 44%, $p = 0.042$, Fisher's exact test). A cue is the visual support that guides the viewer's attention (Ploetzner & Lowe, 2012). Cueing is similar to 'signalling' and can be divided in three classifications of functions for cueing: selection (i.e. cues that guide attention to specific locations, Figure 6.1), organization (i.e. cues emphasize structure, Figure 6.2), and integration (i.e. cues explicate relations between and within elements, Figure 6.3) (De Koning, Tabbers, Rikers, & Paas, 2009). Visual cues such as arrows, might improve students' learning, however guiding attention to specific locations in the display does not guarantee derivation of essential causal relations (De Koning et al., 2009).

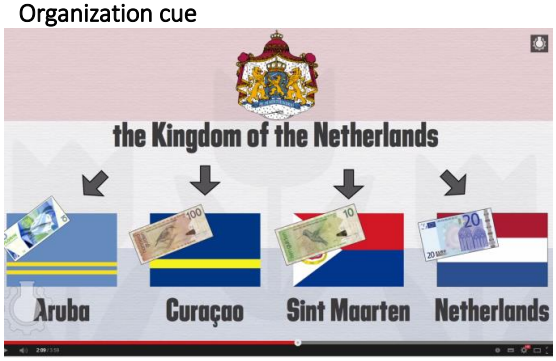
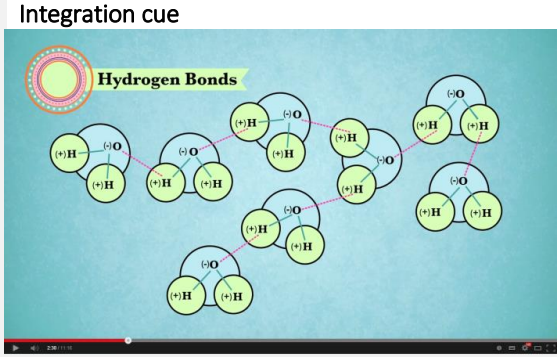
6.1	**The European Union explained** (CGPGrey, 2013) *Narration* So Portugal and Spain have islands from their colonial days that they've never parted with: **these are the Madeira and Canary Islands are off the coast of Africa and the Azores well** into the Atlantic. Because these islands are Spanish and Portuguese they're part of the European Union as well. *Timing* 3.40-3.53 (13 seconds)	Selection cue 
6.2	**Holland vs. the Netherlands** (CGPGrey, 2012) *Narration* The Kingdom of the Netherlands contains three more countries and to find them we must sail from the icy North Sea to the Caribbean and Aruba, Curaçao, Sint Maarten. These are no territories, but self-governing countries within the Kingdom of the Netherlands **and as such, they have their own governments, and their own currencies.** *Timing* 1.55-2.10 (15 seconds) *Signaling* Objects appear synchronic with words.	Organization cue 
6.3	**Water** (CrashCourse, 2012b) *Narration* Essentially what happens is that the positive pole around those hydrogen atoms bonds to the negative pole around the oxygen atoms of a different water molecule. *Timing* 2.25-2.33 (8 seconds) *Signaling* A pink dotted line appears between the molecules	Integration cue 

Figure 6 - Examples of different types of cues

The use of labels differs by popularity, χ^2 (4, N = 71) = 16.282, p = .003. Good videos use more often labels compared to average and poor videos (92% versus 76% and 52%, p = 0.007, Fisher's exact test). Labels are on-screen presented keywords or titles corresponding to the explanation and represent important terms. According to the *modality principle* it is important to reduce on-screen text, how is this matching with the use of labels? Clark and Mayer (2011) make an exception for information the learner needs time to process, such as new terminology. Text presented on the screen as technical or unfamiliar terminology or keywords, can be useful for memory support (Clark & Mayer, 2011). Labels can help to identify entities, to establish relationships, and to organise hierarchies (Ploetzner & Lowe, 2012). The use of labels and names in combination with the narrative and visuals create a memorable image that will stimulate recall, this is called the keyword technique (Smith & Ragan, 2005). Labels shown in the dataset are for example titles and on-screen presented keywords (Figure 6.3 & 9.5). In Figure 7 an example is given of label use. However, in this case an overkill of labels is given at one moment. On-screen text need to be processed and takes cognitive load of the visual channel (Clark &

Mayer, 2011). In addition, the labels are not directly related to the information provided in the audio narration. As a result, the viewer is distracted from the essential information. Cueing could help to improve this video, for example by highlighting the words that are presented in the narration. Although, in this case, removing unused key terms will provide a calm image and will better fit the narrative.

7.1 **The Seed Germination Process** (BackToConstitution, 2012)

Narration
The seed contains a miniature plant, called an embryo that resembled an adult plant, complete with leaves and a route. The seed leaves are called cotyledons. Seeds that contain one embryonic leave, for known as monocotyledons or monocots. **Seeds with two embryonic leaves are termed dicotyledons or dicots.**

Timing
2.25-1.52 (8 seconds)

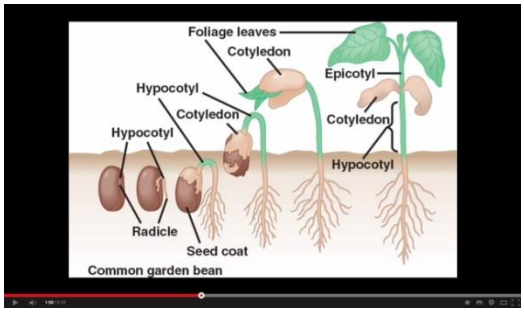


Figure 7 - Example of poor label use

The percentage of spoken prompts differs by popularity, χ^2 (4, N = 71) = 11.244, p = .024. Good videos use spoken prompts more often (70,8% versus 45,8% and 44,5%, p = 0.002, Fisher's exact test). Prompts are questions and requests that intend to facilitate the viewer's cognitive and metacognitive processes (Ploetzner & Lowe, 2012). Prompts can be presented written and spoken (Ploetzner & Lowe, 2012). However, written prompts are less used and show no significantly difference between the popularity groups, which is in line with reducing on-screen text. Two types of spoken prompts are seen, questions and comments. The questions are asked to stimulate the viewer to think about the subject. Questions can be asked at the beginning of the video (Figure 9.1), but also throughout the video (Figure 9.5). The comments are added to help the viewer to become aware of their own learning process. For example, the presenter provides the option to watch parts of the video again (Figure 8), so information is repeated, which is a well-known strategy to increase remembering (Bransford et al., 2000).

Water (CrashCourse, 2012b)

Narration
I think I need some water, but while I'm drinking, **there's a review for all of the things we talked about today.** If there are a couple things you're not quite sure about just go back and watch them. It's not going to take a lot of your time. And you're going to be smarter, I promise.

Timing
10.51-11.07 (16 seconds)

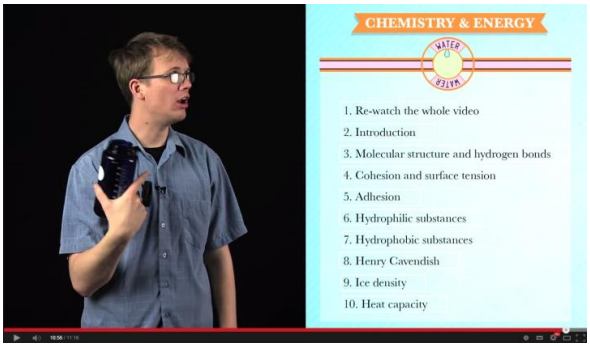


Figure 8 - Example of a spoken prompt presented as comment

Personalisation
No statistical differences were found between personalisation variables and popularity (p > .05). This is in contrast with what is expected, since many studies show that personalisation is

important for videos, for example the use of a conversational style (Clark & Mayer, 2011; Guo et al., 2014; van der Meij & van der Meij, 2013). Although, a clear pattern can be seen in the use of a video-recorded presenter, this difference is just not significant.

Unique characteristics of the three popularity groups

Following the found characteristics, what does a typical good, average and poor video look like? To give an overview, all the important characteristics are summarized in Table 8. Significant characteristics that are presented in more than 60% of the videos are written down as core components of the popularity group. To make this table more concrete, an illustration of a typical good, average and poor video will be given.

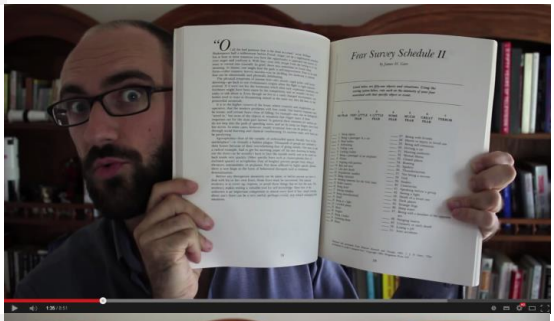
Table 8 - Average (unique) characteristics of the different popularity groups

	Good	Average	Poor
Physical *Representations*	▪ Combination of visual materials with: ○ Pictures (iconic & analytic) ▪ Combination of audio materials with: ○ (Partly) music ▪ Narration in audio ▪ Optional subtitles	▪ Narration in audio ▪ No music ▪ No subtitles	▪ Music ▪ No subtitles, sometimes pertinent
Timing	▪ Around 7 minutes ▪ Slightly quick speaking	▪ Around 5 minutes ▪ Natural speaking ▪ Natural pauses	▪ Around 4 minutes ▪ Natural speaking ▪ Natural pauses
Production quality	▪ HD (720-1080p) ▪ Visual is of high quality ▪ Audio is free of mechanical noise and extraneous audio	▪ 480p	▪ 360-480p ▪ Visual is blurred
Structure	▪ Theoretical explanation with illustrative examples	▪ Illustrative example	▪ Illustrative example
Support	▪ Cues ▪ Spoken prompts ▪ Labels	▪ Labels	
Other	2 years online 60% of video is viewed 200.000 channel members	3 ½ years online 50% of video is viewed 80.000 channel members	3 ½ years online 40% of video is viewed 5.000 channel members

The good video

To illustrate a typical 'good' video, we will look at the top 5 popular educational videos in the summer of 2013 (YouTube, 2013b). The second most popular video corresponds strikingly well with the average characteristics of a good video. On July 2, 2013, YouTube user 'Vsauce' uploaded an instructional video titled, "Why are things creepy?". The video is presented with a high production quality; this is reflected in the HD resolution (1080p). In addition, the audio quality is free of mechanical noise and the video does not contain blurred visuals. The total length of the video is 8 minutes and 50 seconds. The video-recorded presenter asks provocative

questions when he determines a phenomenon (Figure 9.1 & 9.5), which arouses curiosity by the viewer. Stimulating curiosity is an approach to arouse interest and motivation, which is especially relevant when the application of the learning in everyday life lies in the distant future or is unclear (Smith & Ragan, 2005). The presenter illustrates the phenomenon with examples such as pictures, videos, and stories (e.g. Figure 9.1). The music is adapted to the story. While providing examples, the music is mysterious, however when explaining the phenomenon the music is changed into a neutral sound. The explanation contains several perspectives; six different references to research studies, books and academics are given (e.g. Figure 9.2, 9.3 & 9.4). In addition, labels (e.g. Figure 9.5) are used to support the learning process. In short, you will be swept away by the story, but above all, it leads to one goal: to improve the understanding of the topic.

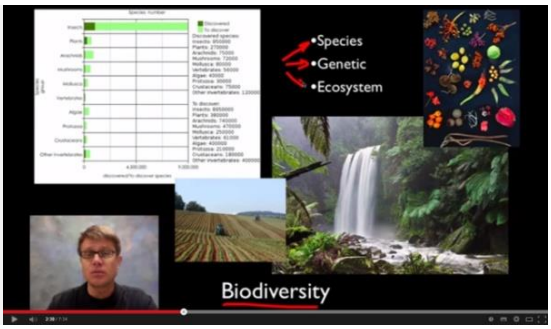
9.1	*Narration* Fear gives us life. Being afraid of the right things kept our ancestors alive. It makes sense to be afraid of poisonous insects or hungry tigers, but what about fear when there is no clear and obvious danger? For instance, **a Teddy Bear with a full set of human teeth.** [...] What gives us the creeps? What causes something to be creepy? *Timing* 0.05-1.10 (65 seconds)	
9.2	*Narration* **Psychologist James Geer developed the "Fear Survey Schedule II"** which he used to find out what scared us the most. *Timing* 1.30-1.40 (10 seconds)	
9.3	*Narration* I love the way **Stephen King delineates three types of scary stuff...** *Timing* 1.54-1.58 (4 seconds)	
9.4	*Narration* For instance, masks, and **why clowns are creepy.** Claude Levi-Strauss wrote that the facial disguise temporarily eliminates, from social intercourse, the part of the body, which reveals personal feelings and attitudes. [...] A mask hides the true emotions and intentions of the person underneath. *Timing* 3.03-3.31 (28 seconds)	

9.5	*Narration* Our language reflects the gray area of terror and creepiness. Take a look at the word 'terror' itself. **We have 'horrible' and 'horrific'. 'Terrible' and …**'terrific'. Why is that? *Timing* 7.15-7.33 (8 seconds)	

Figure 9 – Fragments of the popular video 'Why are things creepy?' created by Stevens (2013)

The average video

In the category average, several videos are uploaded by teachers. Average videos are quite good in explaining the subject; however the video is experienced less fancy. The speed is natural and natural pauses are used. The enthusiasm is missing. This is clearly visible in the following example. On December 30, 2011, a biology teacher uploaded a video about biodiversity as part of a series of videos. The video is presented with an average production quality, which is reflected in the resolution of 360p. The visuals are occasionally pixelate and text is not always legible (e.g. the chart in Figure 10.1). However, the audio quality is free of mechanical noise. The total length of the video is 7 minutes and 34 seconds, which is slightly longer than the average. The video-recorded presenter speaks naturally and uses natural pauses between changes of topic. The explanation of the subject is good, although not as sparkling as the example of the good video. Use of provocative questions would help the viewer to come along in the story. With some adjustments, this can become a good video. The illustrative example used to explain the term 'keystone' in relation with biodiversity, is quite clear (Figure 10.2). Labels and cues are used frequently. Especially selection and organizational cues are used (e.g. Figure 10.1 & 10.2); with arrows and marks structures between terms and the subject matter are clarified. The video is presented quietly, without music or additions. Except for the times when many images are simultaneously displayed.

10.1	*Narration* And when we talk about biodiversity we could be talking about the actual species that we have. We could be talking about the genes that we have. **Or we could be talking about the ecosystems that we have on our planet.** *Timing* 2.21-2.31 (10 seconds)	

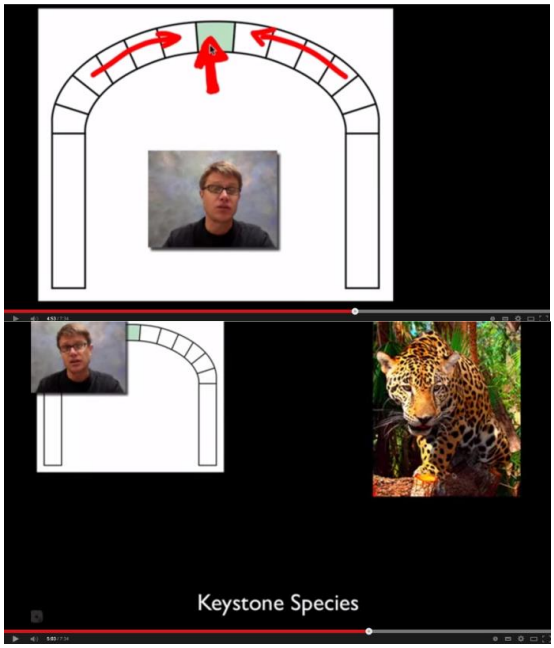
10.2	*Narration* A keystone is a great analogy. So if you're building an arch, an arch essentially has pedestals on either side. You're going to have series of stones that go across the top. But this one block, this one block right up here is called the keystone. [...] **If I were to remove this keystone, all of the weight of this side and all of the weight on this side are all placed against that keystone.** So if I remove that, the whole arch falls in on itself. Now ecosystems are not built like arches but they are to a degree. **And so, a jaguar is an important predator in South America.** *Timing* 4.22-5.03 (41 seconds)	

Figure 10 – Fragments of the average video 'Biodiversity' (Andersen, 2011)

The poor video

To illustrate a poor video is not easy, because it is much harder to find a typical poor video. The differences between poor videos are larger than the differences between good videos. Overall, poor videos are lacking in several characteristics that good videos have. However, which characteristics are lacking, differ per video. This is demonstrated in the following example. A video about the western philosophy presents different parts about its origin and meaning. The information is presented as on-screen text in combination with pictures, however without narration. Several aspects are failing in comparison with the good videos. The production quality of the video is poor. Although the video is uploaded in HD (720p), the visuals are sometimes blurred and texts are many times illegible. The illegible texts are sometimes caused by noisy background pictures (Figure 11.1). The unreadable texts are disastrous, because this video has no narration. It is much harder to perceive the information presented in the video. Another problem is that the visualization is most of the time not in relation with the verbal information (e.g. Figure 11.3). Irrelevant pictures distract the viewer's attention, which can interfere with the process of sense-making (Clark & Mayer, 2011). On the other hand, the quality of the audio is good. The audio is used to present instrumental background music. The length of the video is quite longer than average, with 7 minutes and 32 seconds.

As mentioned before, not all characteristics of good videos are failing in poor videos. In the example video, the structure is made clear in the introduction of the video and the information is summarized in a model (Figure 11.2). In contrast to the majority of poor videos, an illustrative example is missing. However, a theoretical explanation is used, although poorly. According to the topic, you could expect a reference to important philosophers (e.g. Socrates, Plato or Aristotle). However, known data sources and different perspectives are not used. Finally yet importantly, the video contains also supportive characteristics as cues (e.g. arrows) and labels (Figure 11.2).

11.1	*On-screen text* (Translation) Philosophy is a conversation. Especially with yourself. *Timing* 0.55-1.00 (5 seconds)	
11.2	*On-screen text* What we think we see, is the basis for our reasoning. *Timing* 4.11-4.19 (8 seconds)	
11.3	*On-screen text* Philosophy is sceptic and passionate curiosity *Timing* 5.12-5.17 (5 seconds)	

Figure 11 – Fragments of the poor video 'What is philosophy actually?' (Van Delft, 2012)

Guidelines

Following the characteristics of popular conceptual videos (Table 5), an application to the practice can be made through guidelines for video designers. Each characteristic of popular conceptual videos that match with the literature is converted into a guideline. For example, popular videos use audio as narration and provide optional subtitles. Both characteristics are combined in the guideline 'reduce on-screen text' which is in line with the modality principle of Clark and Mayer (2011). The relation of all characteristics with the literature is presented earlier in the discussion. An overview of the guidelines is presented in the conclusion (Table 6).

Conclusion

The purpose of the analysis in this research was to distinguish characteristics of popular and less popular conceptual videos. Many characteristics in the physical dimension differ between the popularity groups. Not the kind of representations matters, but the combination and moreover the production quality of the materials matters. For example, a high resolution, clear audio and sharp visuals. Production quality is a precondition to obtain information from a video, without being distracted. Physical characteristics are also applicable for non-conceptual videos and are well-known from earlier research. Characteristics that are especially relevant for conceptual videos are in the structural and supportive dimension.

The main structural feature of popular conceptual videos is the use of illustrative examples in combination with a theoretical explanation. The main supportive features are the use of labels, cues and spoken prompts. These principles are closely related to instructional design strategies. Smith and Ragan (2005) expand different events of instruction. Two important events in the body of the instruction are 'present information and examples' and 'gain and direct attention'. Presenting information can be done through generalities such as concept definitions which are combined with examples or applications of the principles. To gain and direct attention questions can be asked to help the learner attend to the most critical features of the instruction or by techniques such as zooming, arrows, circles, text boxes etcetera (Smith & Ragan, 2005).

To contribute to the design of conceptual videos, the characteristics of good videos are converted into guidelines (Table 9).

Table 9 - Guidelines for the design of instructional videos for conceptual knowledge development based on the characteristics of good videos

PHYSICAL DIMENSION	
1	**Use a combination of graphics**
1.1	Use pictures and other visual materials to illustrate your story
2	**Reduce on-screen text**
2.1	Use narration rather than on-screen text
2.2	Provide only optional subtitles for non-native speakers
3	**Use high quality materials**
3.1	Record and upload the video in HD
3.2	Use high quality illustrations in the design process
3.3	Prevent mechanical and background noise in the audio
STRUCTURAL DIMENSION	
4	**Use a combination of theory and practice**
4.1	Explain the theory through illustrative examples and stories
SUPPORTIVE DIMENSION	
5	**Support the viewer in their learning process**
5.1	Use labels (Use on-screen keywords)
5.2	Use cues (Use highlighting to guide attention)
5.3	Use spoken prompts (Ask questions to trigger the viewer to search for answers and to reflect)

For the attentive reader: indeed a unique characteristic of popular videos on this list is missing. Namely, the aspect of timing; the longer length of popular videos, with in contrast the

fast speaking and the little use of natural breaks. The reason that this aspect is not included in the guidelines is that further research on this subject is needed, because it is a discussion point in literature. Probably a longer length is not causing popularity, but the attraction of the video results in watching the video for a longer time.

Two limitations in this study are worth naming. At first, the focus of this study is on popular videos. Popularity provides us with information about what people appreciate. However, what people appreciate is not the same as what the best is for learning. Secondly, it is important to take into account that the results of this research are based on videos that maintain certain conditions, such as a minimum number of 1000 views. As a result, extremely poor videos are not included in this research.

Since this research is based on popularity and ratings and not on formal assessments of learning, further research is needed to investigate the learning effects of conceptual videos. Participants can be tested for their knowledge before and after watching the video, and possibly even after a period.

To support new research on video instruction, the used framework is converted into a questionnaire (Appendix D). The questionnaire can be used for the analysis of instructional videos. To improve the framework, several items are transformed to a 7-point Likert scale, so nuances can be distinguished. Furthermore, some characteristics are removed. The characteristics 'symbols 'and 'formal notations' are removed, because the uses of them are depending on the topic of the video and not the popularity. In addition, the characteristic of 'dimensionality' is removed, because this item can be interpreted in different ways for dissimilar graphics, which makes the result unreliable. However, dimensionality could be of influence with evaluating models and animations. Last but not least, it is important to take into account that the perception of the viewer has influence on the outcomes of the framework.

This research has taken a step toward understanding the characteristics of effective instructional videos for conceptual knowledge development. Further research will certainly give rise to additional characteristics as well as to new ways of organising them. The characterisation of instructional videos for conceptual knowledge development makes a basis for analysis of existing videos available. Besides, this study can also inspire the development of new instructional videos as well as the investigation of their educational effectiveness.

References

Andersen, P. (2011). Biodiversity [Video file]. Retrieved from https://www.youtube.com/watch?v=0-PE3ve3w2w

Anderson, L., Krathwohl, D., Airiasian, W., Cruikshank, K., Mayer, R., & Pintrich, P. (2001). A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational outcomes: Complete edition: New York: Longman.

BackToConstitution. (2012). The Seed Germination Process [Video file]. Retrieved from https://www.youtube.com/watch?v=3Ij1eW_gsrM

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education*, 115-131.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn*: National Academy Press Washington, DC.

Bueters, G. (2002). *Handboek voor televisiemakers*: Eburon Uitgeverij BV.

Carliner, S. (2000). Physical, cognitive, and affective: A three-part framework for information design. *Technical communication, 47*(4), 561-576.

CGPGrey. (2012). Holland vs. the Netherlands [Video file]. Retrieved from https://www.youtube.com/watch?v=eE_IUPInEuc

CGPGrey. (2013). The European Union Explained [Video file]. Retrieved from https://www.youtube.com/watch?v=O37yJBFRrfg

Clark, R. C., & Mayer, R. E. (2011). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*: Wiley. com.

CrashCourse. (2012a). Biological Molecules - You Are What You Eat: Crash Course Biology #3 [Video file]. Retrieved from https://www.youtube.com/watch?v=H8WJ2KENlK0

CrashCourse. (2012b). Water - Liquid Awesome: Crash Course Biology #2 [Video file]. Retrieved from https://www.youtube.com/watch?v=HVT3Y3_gHGg

De Haan, J. (2012). Why do competitors do open their stores next to one another? [Video file]. Retrieved from https://www.youtube.com/watch?v=jILgxeNBK_8

De Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review, 21*(2), 113-140.

Fassbender, E., Richards, D., Bilgin, A., Thompson, W. F., & Heiden, W. (2012). VirSchool: The effect of background music and immersive display systems on memory for facts learned in an educational virtual environment. *Computers & Education, 58*(1), 490-500.

Guo, P. J., Kim, J., & Rubin, R. (2014). *How video production affects student engagement: An empirical study of mooc videos.* Paper presented at the Proceedings of the first ACM conference on Learning@ scale conference.

Hiebert, J. (2013). *Conceptual and Procedural Knowledge: The Case of Mathematics*: Taylor & Francis.

Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior, 28*(3), 820-831.

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). *Understanding in-video dropouts and interaction peaks inonline lecture videos.* Paper presented at the Proceedings of the first ACM conference on Learning@ scale conference.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice, 41*(4), 212-218.

Magner, U. I. E., Glogger, I., & Renkl, A. (2014). Which features make illustrations in multimedia learning interesting? *Educational Psychology*(ahead-of-print), 1-18.

Morain, M., & Swarts, J. (2012). YouTutorial: A framework for assessing instructional online video. *Technical communication quarterly, 21*(1), 6-24.

Moran, M., Seaman, J., & Tinti-Kane, H. (2011). Teaching, Learning, and Sharing: How Today's Higher Education Faculty Use Social Media. *Babson Survey Research Group*.

myAlevelBiology. (2013). AS Biology - Enzyme Action [Video file]. Retrieved from https://www.youtube.com/watch?v=Q_ZspsR00kQ

Ploetzner, R., & Lowe, R. (2012). A systematic characterisation of expository animations. *Computers in Human Behavior, 28*(3), 781-794.
Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of educational psychology, 91*(1), 175.
Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of educational psychology, 93*(2), 346.
Ruedlinger, B. (2012). Does Length Matter? It Does For Video: 2K12 Edition. Retrieved from http://wistia.com/blog/does-length-matter-it-does-for-video-2k12-edition
Smith, P. L., & Ragan, T. J. (2005). *Instructional design* (Third ed.): Wiley New York, NY.
Stevens, M. (2013). Why Are Things Creepy? [Video file]. Retrieved from https://www.youtube.com/watch?v=PEikGKDVsCc
Swarts, J. (2012). New Modes of Help: Best Practices for Instructional Video. *Technical Communication, 59*(3), 195-206.
Van Delft, R. (2012). Wat is filosofie eigenlijk? [Video file]. Retrieved from https://www.youtube.com/watch?v=opX_XbqGuWg
van der Meij, H., & van der Meij, J. (2013). Eight Guidelines for the Design of Instructional Videos for Software Training. *Technical Communication, 60*(3), 205-228.
Wember, B. (1976). *Wie informiert das Fernsehen?* : List.
YouTube. (2013a). Statistics Retrieved December, 4, 2013, from http://www.youtube.com/yt/press/statistics.html
YouTube. (2013b). Top 5 Educational Blockbusters of the Summer. from YouTube http://youtube-trends.blogspot.nl/2013/09/top-5-educational-blockbusters-of-summer.html

Appendix

A. Framework for analyzing instructional videos

	Subgroups	Characteristics	Code	ID
Physical Dimension				
Representations	Dynamic graphics	Realistic video	0-3	A
		Animation	0-3	B
	Static illustrations	Iconic pictures	0-3	C
		Analytic pictures	0-3	D
	Text	Symbols	0-3	E
		Formal notations	0-3	F
	Subtitles	None	0	H
		Optional	1	
		Pertinent	2	
	Audio	Narration	0-3	I
		Speech	0-3	J
		Sound	0-3	K
		Music	0-3	L
Timing	Duration	Minutes	00:00	M
	Segmentation	Narration speed is slow	0	N
		Narration speed is natural	1	
		Narration speed is fast	2	
		Natural pauses are included	0-3	O
Production quality	Resolution	144p	0	P
		240p	1	
		360p	2	
		480p	3	
		720p (HD)	4	
		1080p (HD)	5	
	Dimensionality	2D	0	Q
		3D	1	
	Graphics and Text	Blurred	0-3	R
	Audio	Noisy	0-3	S
Cognitive Dimension				
Structure	Intro	Introduction on content	0-3	T
		Goal presented	0-3	U
	Theory	Theoretical explanation	0-3	V
	Practice	Illustrative example	0-3	W
	Evaluation	Summary or conclusion	0-3	X
Coherency	Synchronically	Visual/audio is synchronic	0-3	Y
	Relational	Visual/audio is clearly related	0-3	Z
		Title explain content of video	0-3	AA
	Relevance	Content is relevant for goal	0-3	AB
Extraneous materials	Extraneous elements	Visuals	0-3	AC
		Words	0-3	AD
		Audio	0-3	AE

Affective Dimension					
Scaffolding		Support	Cues	0-3	AF
			Labels	0-3	G*
			Written prompts	0-3	AG
			Spoken prompts	0-3	AH
Personalisation		Style	Formal - Conversational	0-2	AI
		Presenter	No personal visualization	0	AJ
			Video-recorded presenter	1	
			Virtual agent	2	
User control		Navigation functions	Play	**	AK
			Pause		
			Stop		
			Rewind		
		Control functions	Volume		
			Speed		
		Scaling	Zoom		
			Perspective		
			Magnify		

Legend			Coding
Visual	Independent characteristic (Nominal)	0-3	0 = N/a
Auditory	Dependent characteristic (Ordinal)		1 = No
	No classification (Ratio)		2 = Sometimes or Unclear
			3 = Yes

** Removed item*
** User control settings are described in general for YouTube*

B. External recorded data of instructional videos

External data			
Descriptive	Url	Statistical data	Views
	Channel		Likes
	Title		Dislikes
	Subject		Times Shared
	Language		Upload date
			Download date
			Video length
			Average time watched
			Channel members

C. Overview Main Results

Item		Reliability	*Popularity			
ID	*Variable name*	*Weighted Kappa*	*Pearson's Chi-square*			*Fisher's*
			χ^2	*df*	*p*	*Exact test**
Physical Dimension						
Representations						
A	Realistic	1.00	1,430	2	.489	.280
B	Animation	.667	1,000	4	.703	.469
C	Iconic pictures	**-.333**	11,167	4	**.025**	**.003**
D	Analytic pictures	.667	16,875	4	**.002**	**.001**
E	Symbols	.438	4,644	4	.326	.361
F	Formal notations	**.250**	5,644	4	.227	.295
H	Subtitles	1.00	40,005	4	**.000**	-
I	Narration	1.00	1,482	2	.477	.276
J	Speech	.750	2,574	2	.276	.214
K	Sound	.667	3,619	4	.460	.222
L	Music	.667	8,666	4	.070	.106
Timing						
N	Narration speed	**-.087**	26,219	6	**.000**	-
O	Natural pauses are included	.375	18,984	6	**.004**	**.000**
Production quality						
P	Resolution	1.00	38,032	8	**.000**	-
Q	Dimensionality	1.00	10,662	4	**.031**	-
R	Visual Blurred	.667	24,986	6	**.000**	**.011**
S	Noisy Audio	**.000**	8,128	2	**.017***	**.003**
Structural Dimension						
Structure						
T	Introduction on content	.800	2,107	4	.716	.141
U	Goal presented	.667	6,745	4	.150	.600
V	Theoretical explanation	**.250**	4,092	4	.394	.300
W	Illustrative example	**.077**	4,493	4	.343	**.038**
X	Summary or conclusion	.824	4,902	4	.297	.082
Coherency						
Y	Visual/audio is synchronic	.625	4,849	4	.303	.158
Z	Visual/audio is related	1.00	8,434	4	.077	.653
AA	Title explain content of video	.500	5,333	4	.255	.084
AB	Content is relevant for goal	.538	5,506	4	.239	.127
Extraneous materials						
AC	Extraneous Visuals	.625	4,912	4	.296	.226
AD	Extraneous Words	.727	5,189	4	.268	.270
AE	Extraneous Audio	**-.125**	4,554	2	.103	**.027**

Item		Reliability	*Popularity			
ID	*Variable name*	*Weighted Kappa*	*Pearson's Chi-square*			*Fisher's*
			χ^2	*df*	*p*	*Exact test**
Supportive Dimension						
Scaffolding						
AF	Cues	.438	7,108	2	.130	**.042**
G	Labels	.538	16,282	4	**.003**	**.007**
AG	Written prompts	1.00	2,541	4	.637	.253
AH	Spoken prompts	.615	11,244	4	**.024**	**.002**
Personalisation						
AI	Style	**.368**	6,947	4	.139	-
AJ	Presenter (Total)	1.00	7,404	4	.116	-
	Video-recorded presenter	-	3,818	2	.148	.060
	Animated agent	-	4,110	2	.128	.441
Combined Characteristics						
A-D	Type of graphic	-	18,718	7	**.009**	-
	Combination of graphics	-	8,654	2	**.013**	**.003**
H-J	On-screen text instead of narration	-	7,876	2	**.019**	-
I-L	Audio	-	21,773	4	**.000**	-
T-W	Combination Theory/Practice	-	7,498	2	**.024**	-
Other relations						
P*DAYS	Days online*Resolution	-	18,349	4	**.000**	-
P*R	Blurred visuals*Resolution	-	76,441	12	**.000**	-
Fixed data	*(ANOVA-test)*		*F*	*df*	*p*	
M	Minutes	1.00	4,462	1	**.038****	
DAYS	Days online	-	5,494	2	**.006**	-
CM	Channel members	-	24,134	2	**.000**	-
PW	Percentage watched	-	8,567	2	**.003**	

** Answers Yes and Sometimes combined, ** Average and Poor group combined*

Fixed data	*Total* (*N* = 75)		*Good* (*N* = 25)		*Average*(*N* = 25)		*Poor*(*N* = 25)	
	M	SD	M	SD	M	SD	M	SD
Views	961641	1919897	2484266	2592252	165174	223301	235483	985135
Likes	10757	25460	31884	36068	309	538	78	140
Dislikes	326	188	862	950	37	74	81	188
Total Ratings	11084	26011	32746	36776	346	610	159	326
Times shared (*N*= 40)	1708	3084	3229	3664	44	67	11	15
Days online	1142	763	753	721	1355	706	1318	733
Minutes	0:03:50	0:01:45	0:06:55	0:04:38	0:05:07	0:03:30	0:04:42	0:03:27
Percentage watched (*N*= 20)	55,6%	11,9%	60,6%	7,7%	52,5%	1,4%	39,9%	13,8%
Channel members	716590	151961	2068367	2054461	76515	204022	4887	21839

Item			Percentage (%)*			
ID	*Variable name*		*Total* (*N* = 75)	*Good* (*N* = 25)	*Average* (*N* = 25)	*Poor* (*N* = 25)
Physical Dimension						
Representations						
A	Realistic		41,3	**48,0**	44,0	32,0
B	Animation		52,0	**56,0**	44,0	**56,0**
C	Iconic pictures		52,0	**76,0**	36,0	44,0
D	Analytic pictures		37,3	**68,0**	28,0	16,0
E	Symbols		21,3	**22,0**	20,0	12,0
F	Formal notations		21,4	**32,0**	20,0	12,0
H	Subtitles	*No subtitles*	55,4	20,0	**80,0**	66,7
		Optional subtitles	32,4	**76,0**	16,0	4,2
		Pertinent subtitles	12,2	4,0	4,0	**29,2**
I	Narration		62,7	56,0	**72,0**	60,0
J	Speech		32,0	**40,0**	36,0	20,0
K	Sound		21,3	**32,0**	16,0	16,0
L	Music		64,0	**80,0**	40,0	64,0
Timing						
N	Narration speed	*Slow*	7,4	0,0	8,3	**15,0**
		Natural	69,1	50,0	**87,5**	80,0
		Fast	20,6	**50,0**	4,2	5,0
O	Natural pauses are included		76,5	54,2	**95,8**	80,0
Production quality						
P	Resolution	*240p*	13,3	4,0	16,0	**20,0**
		360p	18,7	8,0	12,0	**36,0**
		480p	18,7	4,0	**40,0**	12,0
		720p (HD)	22,7	20,0	**24,0**	**24,0**
		1080p (HD)	26,7	**64,0**	8,0	8,0
Q	Dimensionality	*2D*	56,0	**56,0**	44,0	**56,0**
		Both 2D and 3D	28,0	4,0	8,0	**28,0**
		3D	16,0	40,0	**48,0**	16,0
R	Visual Blurred		46,6	12,0	56,0	**72,0**
S	Noisy Audio		24,6	4,2	**36,0**	33,3
Structural Dimension						
Structure						
T	Introduction on content		57,3	**68,0**	56,0	48,0
U	Goal presented		24,3	24,0	**28,0**	20,8
V	Theoretical explanation		78,4	**84,0**	**84,0**	66.7
W	Illustrative example		68,9	**84,0**	64,0	58,3
X	Summary or conclusion		35,1	**48,0**	24,0	33,3
Coherency						
Y	Visual/audio is synchronic		88,9	**96,0**	91,7	78,3
Z	Visual/audio is related		91,7	**100**	95,8	78,3
AA	Title explain content of video		72,0	**84,0**	60,0	72,0
AB	Content is relevant for goal		88,0	**96,0**	88,0	80,0

Item			Percentage (%)*			
ID	*Variable name*		*Total* (*N* = 75)	*Good* (*N* = 25)	*Average* (*N* = 25)	*Poor* (*N* = 25)
Extraneous materials						
AC	Extraneous Visuals		22,7	24,0	12,0	**32,0**
AD	Extraneous Words		13,5	8,0	**20,0**	12,5
AE	Extraneous Audio		17,8	4,2	24,0	**25,0**
Supportive Dimension						
Scaffolding						
AF	Cues		52,0	**68,0**	52,0	36,0
G	Labels		73,4	**92,0**	76,0	52,0
AG	Written prompts		25,3	**32,0**	28,0	16,0
AH	Spoken prompts		54,6	**70,8**	45,8	44,5
Personalisation						
AI	Style	*Formal*	28,4	20,0	24,0	**41,7**
		Both formal and conversational	60,8	**76,0**	56,0	50,0
		Conversational	10,8	4,0	**20,0**	8,3
AJ	Presenter	*None*	70,7	**40,0**	24,0	24,0
		Video-recorded presenter	26,7	**40,0**	24,0	16,0
		Animated agent	2,7	0,0	0,0	4,0
Combined Characteristics						
A-D	Combination of graphics		52,0	**76,0**	40,0	40,0
H-J	On-screen text instead of narration		10,7	4,0	4,0	**24,0**
I-L	Audio	*Narration or speech*	29,3	8,0	**52,0**	28,0
		Music or sounds	12,0	4,0	4,0	**28,0**
		Both	58,7	**88,0**	44,0	44,0
T-W	Combination Theory/Practice		51,4	**72,0**	48,0	33,3

** NB: The percentages in the structural dimension reflect only the times answering 'yes' and not 'unclear'.*

D. Questionnaire for analyzing instructional videos [New]

ID: _____ Title: _____

Physical Dimension

Which visual representations occur in the video?

- □ Realistic video
- □ Animation
- □ Pictures
  - o Iconic pictures
  - o Analytic pictures

Are subtitles used in the video?

- □ None
- □ Optional
- □ Pertinent

Which kind of audio materials are used?

- □ Narration
- □ Speech
- □ Sound
- □ Music

What is the length of the video?

_ _ : _ _ : _ _ minutes

	N/a	Slow	Slightly slow	Natural	Slightly fast	Fast
The narrations speed is	0	1	2	3	4	5

	N/a	Never	Very rarely	Rarely	Occasio-nally	Frequently	Very frequently
Natural pauses are used in the video	0	1	2	3	4	5	6

The resolution of the video is:

- □ 144p
- □ 240p
- □ 360p
- □ 480p
- □ 720p (HD)
- □ 1080p (HD)

	N/a	Clear/Bright					Blurred
Graphics in the video are	0	1	2	3	4	5	6

	N/a	Clear/bright					Noisy
Audio in the video is	0	1	2	3	4	5	6

Structural Dimension

	N/a	Totally Disagree					Totally Agree
An introduction is given	-	1	2	3	4	5	6
A goal is presented	-	1	2	3	4	5	6
A theoretical explanation is given	-	1	2	3	4	5	6
A illustrative example is used	-	1	2	3	4	5	6
A summary or conclusion is given	-	1	2	3	4	5	6
Visual and audio in the video are synchronic	0	1	2	3	4	5	6
Visual and audio is clearly related	0	1	2	3	4	5	6
The title explain the content of the video	-	1	2	3	4	5	6
Content is relevant for goal	-	1	2	3	4	5	6
Extraneous visuals occur	0	1	2	3	4	5	6
Extraneous words are used	0	1	2	3	4	5	6
Extraneous audio occur	0	1	2	3	4	5	6

Supportive Dimension

	N/a	Never	Very rarely	Rarely	Occasio-nally	Frequently	Very frequently
Cues are used	-	1	2	3	4	5	6
Written prompts are used	-	1	2	3	4	5	6
Spoken prompts are used	0	1	2	3	4	5	6
Labels are used	0	1	2	3	4	5	6

	N/a	Formal					Conversational
The narration style of the video is	0	1	2	3	4	5	6

A personal visualization is presented

- ☐ No personal visualization
- ☐ Video-recorded presenter
- ☐ Virtual agent

Which of the following user control functions are available?

Navigation functions

- ☐ Play
- ☐ Pause
- ☐ Stop
- ☐ Rewind

Control functions

- ☐ Volume control
- ☐ Hide audio
- ☐ Speed control

Scaling functions

- ☐ Zoom
- ☐ Perspective
- ☐ Magnify