

A Conceptual Characterization of Online Videos Explaining Natural Selection

Gustav Bohlin¹ · Andreas Göransson¹ ·
Gunnar E. Höst¹ · Lena A. E. Tibell¹

Published online: 2 December 2017

© The Author(s) 2017. This article is an open access publication

Abstract Educational videos on the Internet comprise a vast and highly diverse source of information. Online search engines facilitate access to numerous videos claiming to explain natural selection, but little is known about the degree to which the video content match key evolutionary content identified as important in evolution education research. In this study, we therefore analyzed the content of 60 videos accessed through the Internet, using a criteria catalog with 38 operationalized variables derived from research literature. The variables were sorted into four categories: (a) key concepts (e.g. limited resources and inherited variation), (b) threshold concepts (abstract concepts with a transforming and integrative function), (c) misconceptions (e.g. that evolution is driven by need), and (d) organismal context (e.g. animal or plant). The results indicate that some concepts are frequently communicated, and certain taxa are commonly used to illustrate concepts, while others are seldom included. In addition, evolutionary phenomena at small temporal and spatial scales, such as subcellular processes, are rarely covered. Rather, the focus is on population-level events over time scales spanning years or longer. This is consistent with an observed lack of explanations regarding how randomly occurring mutations provide the basis for variation (and thus natural selection). The findings imply, among other things, that some components of natural selection warrant far more attention in biology teaching and science education research.

1 Background

1.1 Introduction

Evolutionary theory is one of the most important theories in science (Dobzhansky 1973; Futuyma 1995) given that it underpins all biological disciplines by accounting for the phylogenetic, structural, and functional relatedness of all living things. In addition, it has

✉ Lena A. E. Tibell
lena.tibell@liu.se

¹ Department of Science and Technology, Linköping University, 60174 Norrköping, Sweden

profound implications for complex societal issues such as bacterial resistance to antibiotics and adaptations to climate change. Hence, evolution is clearly of critical importance for biology education. However, learning difficulties and robust misconceptions related to evolution through natural selection are frequent among learners as well as teachers (e.g. Andrews et al. 2012; Gregory 2009; Hiatt et al. 2013; Smith 2010a, b). Although a number of teaching strategies have been evaluated during the last 30 years (Smith 2010b; Understanding Evolution 2017), no reliable solution or “magic bullet” for effective teaching of natural selection has emerged as yet.

The use of visual representations for learning science has many beneficial aspects (Gilbert 2005; Phillips et al. 2010). These include general aspects of learning with visual representations, such as the multimedia principle, which suggests that people learn better from words and pictures in combination than from words alone (Mayer 2001; Fletcher and Tobias 2005). But also specific features of individual visualizations can aid learning, for example using highlighting to emphasize features of interest in an object, or representing a process, pattern, or phenomenon schematically to allow an overview such as a phylogenetic tree or a Punnett square (Phillips et al. 2010). In addition, dynamic visualizations such as videos can be used to efficiently communicate aspects that are difficult to represent using static pictures or diagrams. Examples include videos that can improve students’ understanding of molecular movements and dynamics (Pallant and Tinker 2004) by visualizing *spatial changes* in both position and shape (Rundgren and Tibell 2010). Visualizations can also help learners conceive events and processes that are difficult to grasp because they span temporal and spatial scales far larger or smaller than those that we can directly perceive. Thus, visual representations could make important contributions to fostering understanding of evolutionary processes. In this regard, the relatively scarce attention in educational research on biological visualizations and simulations that is paid to evolution has prompted Lee and Tsai (2013) to recommend further research to address the potential of learning abstract and complex biological concepts through dynamic visualizations. Given the challenges of learning natural selection and the accompanying increase in the number of available online explanatory videos, it is important to examine how emerging digital teaching and learning tools convey the subject.

1.2 Analytical Categories

The following sections will present the underpinnings for the variables utilized in the study. An important conceptual background is found in a framework developed by Tibell and Harms (2017). This framework proposes two main conceptual dimensions that address concepts that apply specifically to biological processes and concepts that relate to phenomena that transcend biology, respectively. Tibell and Harms argue that natural selection consists of three main principles: variation, inheritance, and selection (cf. Godfrey-Smith 2007). At a more fine-grained level, each of these principles is composed of so-called key concepts, which make up the first dimension of the framework. The second dimension consists of concepts of a more abstract nature called threshold concepts, such as randomness and temporal scale. Based on this framework, a potential strategy for supporting learners may be to actively consider the relationship between threshold concepts and key concepts during instruction. In this regard, Tibell and Harms (2017) argue that the use of visualizations is a promising strategy to help learners perceive threshold concepts and how they are related to natural selection.

In addition to exploring the inclusion of key concepts and threshold concepts, we explore the extent to which videos present faulty explanations that correspond to known

misconceptions. Misconceptions are erroneous ideas associated with learning natural selection, such as the belief that evolution is driven by need, or that only beneficial traits are inherited. They can be major obstacles to learning and are frequently expressed by students at all educational levels (e.g. Gregory 2009). Lastly, we study the kinds of organisms that are utilized for explaining evolution in the videos. This has previously been linked to so-called surface feature effects in that different organism contexts tend to invoke different kinds of explanations and misconceptions in novice learners (e.g. Nehm and Ridgway 2011). In summary, the variables used in the study are sorted under the four analytical categories key concepts, threshold concepts, misconceptions, and organismal context, described subsequently.

1.2.1 Key Concepts

Natural selection is often defined in terms of key concepts (e.g. Anderson et al. 2002; Nehm and Reilly 2007) based on Mayr's (1982) description of natural selection as a logical outcome of five premises and three inferences. For example, Anderson et al. (2002) based the Concept Inventory of Natural Selection (CINS) for testing evolutionary knowledge on key concepts and additional concepts. Key concepts of natural selection are also applied in a number of other test instruments, for example those presented by Bishop and Anderson (1990), Nehm and Reilly (2007), and Nieswandt and Bellomo (2009). The key concepts seem to be the most common conceptual framework for studying students' learning of natural selection, and studies typically focus on their understanding and/or misunderstanding of these concepts.

Natural selection will only lead to evolution if there is heritable variation in a population. Such variation may be manifested as morphological, physiological, or behavioral differences among members of a population. While variation in many characteristics has both environmental and genetic origins, the ultimate sources of heritable variation are random mutations and genetic recombination. This leads to individual variation (including variation in reproductive capacity), which is a fundamental requirement for natural selection. Mutations can have a range of effects on fitness (zero, positive, or negative) in relation to the surrounding environment. Fitness is usually regarded as a relative concept defined in terms of factors that have an effect on reproduction and survival (Lewontin 1970). The process of natural selection requires individuals to be able to reproduce and pass genes from generation to generation. Hence, traits that correlate with an increased likelihood to survive and reproduce in that environment (i.e. an increased fitness) tend to accumulate in subsequent generations of the population (Anderson et al. 2002; Mayr 2001; Nehm and Reilly 2007; Nehm and Ha 2011). This process of natural selection may involve many generations, and favorable characteristics may gradually dominate a population.

1.2.2 Threshold Concepts

The last decade has seen an increasing interest in threshold concepts and their postulated importance for learning progression (Meyer and Land 2003, 2005). Ross and collaborators (2010) have identified potential threshold concepts in biology and have argued that such threshold concepts concern processes and abstract ideas are fundamental to biological thinking (Ross et al. 2010; Taylor 2006). Furthermore, biological threshold concepts represent areas that are generally left tacit as "assumed understanding" in teaching (Davies 2006; Ross et al. 2010). This study will focus on four threshold concepts included in the framework developed by

Tibell and Harms (2017): *temporal scale, spatial scale, randomness, and probability*. Their relevance for natural selection is discussed next.

Evolution includes phenomena and mechanisms that span a wide range of *space* and *time*. For example, the geographic range of organisms varies by as much as twelve orders of magnitude (Brown et al. 1996). Evolution involves multiple linked processes, ranging from subcellular events occurring on scales of nanoseconds and nanometers to gradually unfolding developments over millions of years at a global geographical scale (van Dijk and Reydon 2010). Evolutionary processes transcend levels of organization from the origin of variation at the micro-scale to variation at the level of individual organisms and changes in large populations. As a consequence, natural selection concepts might appear disconnected. Understanding deep time in itself has proven difficult (Cheek 2010), and the ability to connect different levels of organization in space and time is challenging for students (Johnstone 1982; Jördens et al. 2016; Lewis and Kattmann 2004). Another crucial aspect of natural selection is the role of random events (e.g. Garvin-Doxas and Klymkowsky 2008; Robson and Burns 2011). Mutations are random with respect to fitness in a given environment; the environment does not selectively favor or induce mutations with positive fitness effects (Morris and Lundberg 2011). However, due to factors that give rise to differential survival within populations, traits that correlate with higher fitness than others are strongly (and non-randomly) favored. The fourth threshold concept in the present study is *probability*, which is closely related to randomness. For example, it is possible to calculate the probability that certain mutations and associated traits will arise randomly in a population of a given size in a given number of generations. Although highly unlikely to arise in any specific individual in the population, the probability that the same mutation will occur eventually in a large population over a sufficient number of generations is very high. Probability is thereby associated to key concepts such as differential fitness or inheritance.

1.2.3 Misconceptions

Teachers, parents, and popular science sources such as television and movies occasionally present incorrect explanations of natural selection (Gregory 2009). This can lead to misconceptions (also referred to as alternative conceptions, see e.g. Leonard et al. 2014) about the underlying mechanisms. Misconceptions can also be referred to as non-normative ideas or naïve ideas derived from cognitive constraints on our thinking about the world (Sinatra et al. 2008; Smith 2010b). Such ideas can be a constraint that hinders change to a more scientific conception of evolution (Smith 2010b). The ideas can persist into adulthood, and research has shown that resulting misconceptions are common even among biology majors (Gregory 2009; Smith 2010b). Major misconceptions related to learning natural selection include essentialism, “use and disuse,” soft inheritance, teleology, and anthropomorphism (Nehm and Schonfeld 2008; Gregory 2009; Bishop and Anderson 1990). Essentialism is the idea that organisms belong to discrete and uniform units that share a hidden causal power or “essence,” an idea which rejects or downplays variation within populations or species. This has been found to be an obstacle for accepting and understanding the concept that species can be subject to constant change and also to the speciation process (Shtulman 2006; Shtulman and Schulz 2008). Darwin is commonly credited for introducing the contrasting idea of variation among individuals within species and populations. However, this shift in evolutionary thought was more complex, and a plurality of views was present both before and after Darwin (see e.g. Kampourakis 2015a, b). “Use and disuse” refers to the incorrect idea that evolution results

from use and/or disuse of organs. Soft inheritance is the erroneous idea that acquired characteristics are inherited, which would mean that changes occurring in individuals have evolutionary consequences. In addition, teleological and anthropomorphic language is often used by experts, teachers, and documentary television programs (Aldridge and Dingwall 2003; Dingwall and Aldridge 2006) as it provides a shortcut for more elaborate and complex explanations.

1.2.4 Organismal Context

Previous research has documented effects of item features in assessment of evolution knowledge (Nehm and Ha 2011; Olson 2012). One of the factors affecting the explanatory pattern is the type of organism used in assessment questions (Nehm and Ha 2011; Opfer et al. 2012; Spiegel et al. 2006). Spiegel et al. (2006) found that insects, diatoms, and viruses were more likely than humans, finches, or whales to elicit intuitive modes of explanation. Potential effects on students' explanations from items involving plants or fungi seem to be absent in the literature. Interestingly, students tend to use different explanations for evolutionary phenomena across different species and for different processes, where the reasons for this seem to remain unclear. In fact, not only do students use different key concepts to describe evolutionary processes in different species, they also seem to adhere to different misconceptions depending on the organism included (Nehm and Ridgway 2011; Opfer et al. 2012; Spiegel et al. 2006). Whether this is due to a limitation in the range of evolutionary examples used in education or due to cognitive factors is an open question. Therefore, it might be of interest to document the range of biological taxa used as examples for natural selection in explanatory sources such as videos or books.

1.3 Research Questions

Three specific research questions were addressed in the study:

- 1) What content with regard to specified (a) key concepts, (b) threshold concepts, and (c) misconceptions associated with natural selection is conveyed by videos on the internet purporting to explain evolution?
- 2) What organismal contexts are used to illustrate natural selection in videos on the internet?
- 3) What patterns with regard to co-occurrence of the identified content can be distinguished in the sampled videos?

2 Method

2.1 Sample

Data from publicly accessible video files were collected using an approach intended to approximate frequently used search strategies. Thus, the sample was compiled via web searches using Google Video and YouTube in combination with searches on the resulting webpages (cf. Ortiz-Cordova et al. 2015). For example, a search using a generic search engine (e.g. Google Video) may yield a result from a platform with educational resources (e.g. Khan academy). In such cases, the search functionality on the web page was employed to search for

additional material on the platform. In addition, certain streaming sites have a functionality where “related videos” appear adjacent to the video currently being played. In these cases, suggested videos were also considered for inclusion (e.g. Lei et al. 2015). The search process was complemented with targeted searches on specific platforms that did not emerge in the main search procedure (ed.ted.com, evolutionfilmfestival.org, Cassiopeia Project, HHMI Biointeractive, Annenberg Learner and Encyclopedia Britannica). Searches were performed using a variety of search terms, including “evolution” and “natural selection” in combination with “video,” “animation,” “visualization,” and “explanation.” This hybrid search strategy was flexible and explorative in the sense of employing multiple search approaches rather than adhering to a strict search term output. Following an initial viewing of identified videos, the videos that conformed to the following criteria were compiled in a database: (1) were about biological evolution (excluding videos that show non-biological evolution e.g. of transistors); (2) were not explicitly arguing against evolution (thus excluding videos with a clear creationist agenda); (3) had an explanatory approach toward evolution. The third criterion was applied generously, including not only videos with titles such as “how evolution works” or “evolution explained” but also more neutral titles that used narration or images to convey information on the workings of evolution. The final sample consisted of 60 videos.

2.2 Coding Scheme

The research questions were addressed through content analysis (Krippendorff 2004). A criteria catalog covering the focal content areas was developed as outlined in the following section, resulting in a grid with 38 criteria for use as variables to capture video content (Table 1). The operationalization, relative frequency of occurrence in the videos, and reliability statistics are presented for each variable in Appendix A. Note that this design allowed us to record the presence or absence of the included variables, but not to assess the quality of the videos (which is clearly of interest, but beyond the scope of this study).

Each variable was operationalized in a dichotomous manner (present or non-present). This design allowed us to explore the presence of pre-defined variables in the videos. Both visual and auditory representational modes were considered, i.e. a variable was coded as present if it was represented either visually or audibly (or both). To prevent subjective interpretation of implicit content, only explicitly presented information was considered. Explicit mentioning of intended audience for the videos as well as video length was also noted.

2.2.1 Generation of Variables

The variables for the content analysis were developed and sorted across four main categories: key concepts, threshold concepts, misconceptions, and organismal context. Moreover, key concepts were grouped under the three main principles suggested by Tibell and Harms (2017): variation, inheritance, and selection. The included key concepts are *origin of variation*, *individual variation*, *inherited variation*, *differential fitness* (including reproductive success), *limited survival*, *change in population*, and *speciation*. Selection pressure was not included as an individual variable due to its problematic association to force-talk and evolutionary misunderstandings (Depew 2013; Nehm et al. 2010). However, selection pressure is largely captured, at least implicitly, by the *limited survival* and *limited resources* variables. The principles (variation, inheritance and selection) were not operationalized into separate variables. Instead, the presence of each of these three principles was analyzed deductively from the key concepts, as described in Section 2.3.

Table 1 Variables in the criteria catalog

KEY CONCEPTS	THRESHOLD CONCEPTS	MISCONCEPTIONS
Variation	Spatial scale	28 Need
1 Origin of variation	12 Molecule	29 Anthropomorphism
2 Individual variation	13 Genome	30 Only good traits are inherited
3 Differential fitness	14 Individual	31 Natural selection as an event
	15 Population	32 Acquired traits are inherited
Inheritance	16 Species	33 Essentialism
4 Inherited variation		
Selection	Transition between organizational levels	ORGANISMAL CONTEXT
5 Limited resources	17 Gene-protein	34 Bacterial evolution
6 Limited survival	18 Protein-cellular function	35 Animal evolution
7 Change in population	19 Cellular function - individual characteristics	36 Human evolution
8 Speciation		37 Plant evolution
		38 Symbolic evolution
Additional variables	Temporal scale	
9 Number of traits	20 Shorter than seconds	
10 Negative mutations	21 Seconds	
11 Extinction	22 Minutes/hours	
	23 Days	
	24 Years	
	25 Generations	
	Randomness	
	26 Randomness	
	Probability	
	27 Probability	

In addition to the abovementioned variables that belong to the three overarching principles, we included three more specific content-related variables corresponding to key concepts. The first was the *number of traits*, which was included to record whether the videos illustrated evolution of one trait or more than one trait. The motivation for monitoring this aspect is the potential risk that showing only one mutation or trait might promote the misunderstanding that only one trait varies at a time. The second was *negative mutations*, i.e. whether or not the videos convey the message that mutations are not always beneficial and may lead to reductions in evolutionary fitness (Gregory 2009). The third was *extinction*, i.e. whether or not loss of species is represented (Shtulman 2006). Although the literature does not suggest that grasping these three concepts is as crucial for understanding natural selection as grasping the other key concepts, they were grouped among the key concepts here given their clear subject-specific relations to the biological content.

Regarding the threshold concepts, *temporal scale* and *spatial scale* were differentiated into several variables, each intended to capture the presence or absence of a specific level of organization (e.g. indications of time in seconds, or spatial events on a population level) in the videos. This approach allowed us to precisely quantify which levels were represented in both the individual videos and the whole set. Regarding time, the variables consisted of *shorter than seconds* and *seconds* (mainly relevant for intracellular processes such as mutations), *minutes/hours* and *days* (of relevance for organisms with short generation times such as bacteria), *years*, and *generations* (to be able to contrast the use of absolute and relative time scales). Regarding spatial scale, we included variables for five levels of organization. In addition, we constructed three variables intended to measure whether the videos made transitions between three defined levels of organization: from *gene to protein*, from *protein to cellular function*, and from *cellular function to individual characteristics*. *Randomness* and *probability* were each assigned one variable.

We included variables associated with some of the common misconceptions to capture whether they are mistakenly promoted by the videos. Since several misconceptions seem to involve confusion regarding the suggested threshold concepts (for example, the belief that evolution is need-driven implies that the role of random factors is not understood), we believe that their possible co-existence with the identified key and threshold concepts warrants attention. The misconceptions targeted in the study were: *need* (that variation appears in response to a need in the environment), *anthropomorphism* (that entities behave as if they have human intentions and ability to plan ahead), *only good traits are inherited* (that all individuals that do not have a specific beneficial trait die before reproducing), *natural selection as an event* (that major changes occur instantaneously rather than through gradual shifts over time), *acquired traits are inherited* (that traits developed during a lifetime are passed on to offspring), and *essentialism* (that when a species changes the essential nature of all of its members changes simultaneously).

Finally, with regard to organismal context, we included five variables to record the kinds of organism(s) used as examples in the videos, based on previous findings: *bacteria*, *animals* (excluding humans), *humans*, *plants*, and *symbolic* (fictional) organisms.

2.3 Analysis

Sixty videos were coded in total. Two coders independently analyzed a subset of the videos with an overlapping sample of 33 videos that were the same for both coders. This sample was used to calculate inter-coder reliability by calculating Krippendorff's alpha coefficients for each variable (Hayes and Krippendorff 2007), using a cut-off value of .70 for acceptable reliability. These coefficients were calculated with IBM SPSS Statistics (version 21) using a macro described by Hayes and Krippendorff (2007). When coders disagreed on whether a variable was present in a video or not, it was coded as present because it is more likely for a coder to miss a variable that is present than to mistakenly see or hear an absent variable. Nevertheless, this may lead to a certain overestimation regarding the presence of some variables that should be acknowledged when interpreting the data.

The coding procedure was followed by analyzing the distribution of concepts within categories. The presence of the three principles (variation, inheritance and selection) in the videos was analyzed indirectly by comparing the inclusion of the key concepts associated with each principle. We adopted two approaches: a relaxed approach to reveal any indications of presence (A), in which the principle was regarded as present if any of the associated key concepts were included; and a strict approach to assess the degree of full representation (B), in which the principle was only regarded as present if all the associated key concepts were

present. Inclusion of temporal scales was analyzed with respect to relative time (i.e. generations) and absolute time.

Patterns and relationships within the data were explored with cluster analysis (Everitt et al. 2011), which consists of methods for revealing groupings of objects or variables based on their similarities. The analysis consisted of two parts, one in which possible relations between variables were investigated and another in which groupings of videos were analyzed (cf. Islam et al. 2014). As a first step, variables to include in the calculations were selected. Here, we excluded 16 variables for which 90% or more of the videos were the same (i.e. variables that were present in $\leq 10\%$ or $\geq 90\%$ of the videos were excluded from analysis, see Appendix A). Such variables contribute little information about differences within the data set (cf. Hamid et al. 2010) and may even mask patterns in the data (Everitt et al. 2011).

When clustering variables, information about correlations is typically of interest (Anderberg 2014). We used Yule's Q, a measure of association between binary variables, as the similarity measure for the variable clustering procedure (e.g. Hamid et al. 2010). Agglomerative hierarchical clustering based on group average linkage (Everitt et al. 2011) was performed. The algorithm corresponds to creating a series of successively larger clusters, starting with the individual variables as initial clusters. New clusters are formed iteratively between the two most similar clusters as determined by the similarity measure, until only a single cluster remains. The

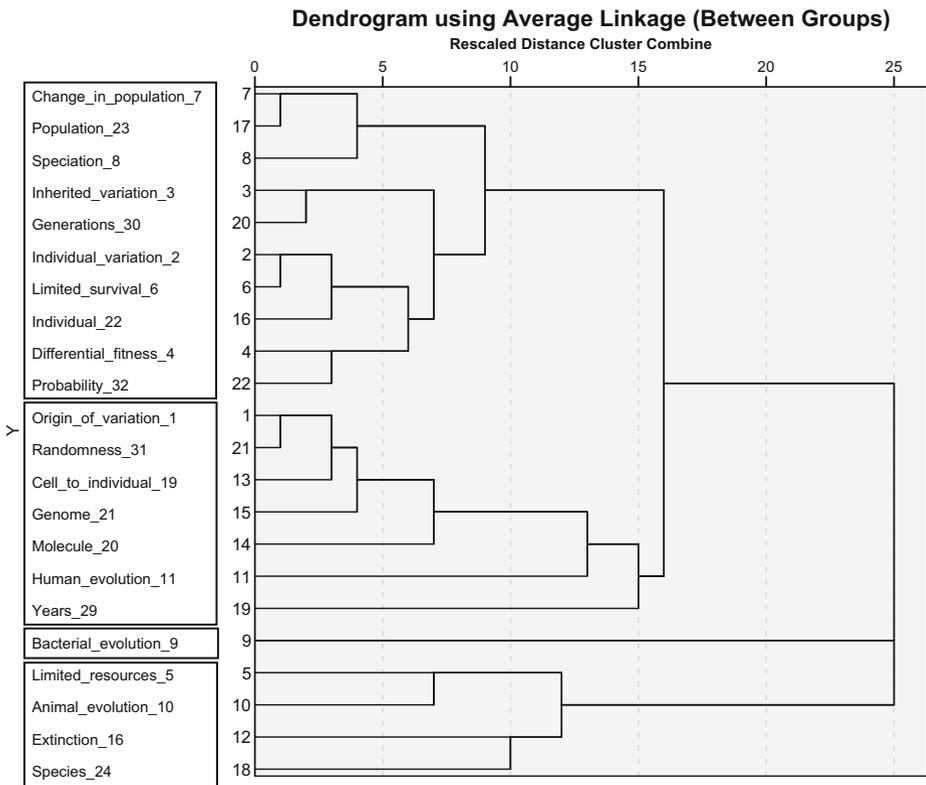


Fig. 1 A dendrogram that illustrates the fusion of clusters during the hierarchical agglomerative clustering procedure. The horizontal axis indicates the distances at which clusters are combined. The four resulting variable clusters are indicated by rectangles

process yielded the dendrogram in Fig. 1. The number of clusters was determined from the agglomerative coefficient, where an “elbow” identified in a plot of the coefficients indicates a point where fusing clusters would mean combining highly dissimilar clusters into one.

A hierarchical agglomerative clustering of videos based on group average linkage was performed using a “matching coefficient” as a measure of similarity. This coefficient is a suitable similarity measure for data where presence as well as absence of a feature is considered informative (Everitt et al. 2011). Given the purpose of the current study, we consider videos that lack a certain variable as sharing a similarity in the same sense as videos that contain the variable. The number of clusters was determined from the agglomerative coefficient as described above.

Lastly, the video-length and explicitly stated audiences in the resulting video-clusters were compared to reveal possible differences.

2.4 Reliability Limitations

Most variables reached the specified reliability threshold (see Appendix A). One reason for the lower Krippendorff’s alpha coefficients for some variables is that they may be more abstract and therefore more difficult to discern with precision. Nevertheless, findings pertaining to the variables with lower reliability should be regarded more cautiously than those regarding other variables. Krippendorff’s alpha requires a higher level of agreement for variables that are present in either a high or a low fraction of videos to reach a value above the cut-off (> 0.70). Therefore, in addition to the reliability value, both the magnitude of the disagreements as well as the relative proportion of disagreements need to be considered in order to establish a meaningful assessment of reliability (Milne and Adler 1999). As an example, a seemingly high agreement of 97% (where 32 of 33 decisions were identical for the variable *time in days*) could still yield an alpha-value of zero. Thus, although results regarding these variables should be interpreted with caution, we nevertheless believe that it is valuable to present these data.

3 Results

The analysis revealed that the majority of videos conveyed few key concepts of natural selection. Among the videos that contained several variables, a typical video from the sample conveyed that the chance of an individual animal to survive and reproduce in a specific environment depends on its traits. Each of these three variables (i.e. spatial scale: Individual; organismal context: Animal; key concept: Differential fitness) was present in 58–75% of all the videos. Meanwhile, few videos showed the events following a mutation in a gene and its expression, and no video explicitly indicated that processes and events with evolutionary relevance might occur in time frames ranging from seconds to hours. The overall distribution of variables is shown in Fig. 2, whereas the relative frequency of occurrence and the corresponding reliability value for each variable are shown in Appendix A. In the following sections, the results for each category of variables are described first, followed by results concerning relations between categories of variables.

3.1 Key Concepts

The most frequently illustrated principle was variation (Table 2). At least one key concept classified under this principle was present in 72% of the analyzed videos, and all three were present in 15% of the videos. Of this 72%, *individual variation* (variable 2, 48%) and

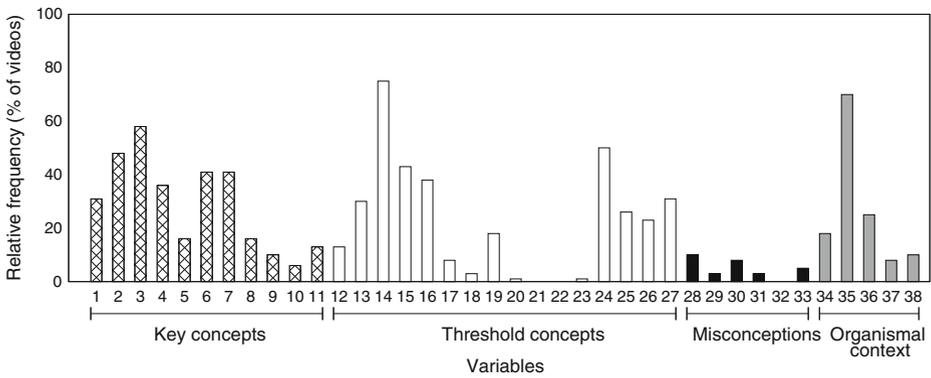


Fig. 2 Relative frequency (%) of the 38 variables (see Table 1) present in the videos ($n = 60$)

differential fitness (variable 3, 58%) were most frequently expressed, while *origin of variation* (variable 1, 32%) was less common. The inheritance principle, in the form of the key concept *Inherited variation* (variable 4), was included in 37% of the videos. The selection principle was represented by at least one key concept in 60% of the videos, but only 5% included all four key concepts (Table 2). The most commonly included key concepts of the selection principle were *limited survival* (variable 6, 42%) and *change in population* (variable 7, 42%). Limited survival was more commonly represented than *limited resources* (variable 5, 17%), indicating that most videos used predation (or no explanation), rather than starvation, for example, to illustrate the notion that not all individuals survive to reproduce. The most frequent combination of principles was the variation and the selection principles, which were simultaneously represented by at least one key concept from each principle in more than half of the videos. Interestingly, 12 (20%) of the 60 videos did not include any key concepts related to the important principles (variables 1–8, Table 1), and only one video conveyed all of them (Table 2). Among the additional content-related key concepts (variables 9–11), *extinction* and *number of traits* were conveyed by 10% of the videos while *negative mutations* were included in four videos (7% of the sample).

3.2 Threshold Concepts

Variables associated with the threshold concepts were conveyed to very different extents across the videos. This was especially apparent for levels of organization in space and time (Fig. 2;

Table 2 Absolute and relative frequency of occurrence of evolutionary principles in the videos, as recorded if (A) at least one and (B) all associated key concepts were present, respectively

Principles	A. At least one key concept linked to each principle was present	B. All key concepts linked to each principle were present
Variation (variables 1–3)	43 (72%)	9 (15%)
Inheritance (variable 4)	22 (37%)	22 (37%)
Selection (variables 5–8)	36 (60%)	3 (5%)
Variation + inheritance	21 (35%)	8 (13%)
Variation + selection	32 (53%)	1 (2%)
Inheritance + selection	19 (32%)	2 (3%)
Variation + inheritance + selection	19 (32%)	1 (2%)

Appendix A). Most videos showed phenomena at the individual level (75%); the population level (43%) was also fairly frequently depicted (note that for events at the individual level, reliability fell below the agreed cutoff level). In contrast, the videos showed genomic or molecular level events (30 and 13%, respectively) less often. The same pattern was even more pronounced for temporal scales, where phenomena were almost exclusively conveyed in years or generations (most often in terms of years) and hardly ever in smaller units, with days and smaller than seconds present in only one video each. *Probability* was explicitly conveyed in 31% of the videos and *randomness* in 23%. These two concepts were usually only communicated orally (results not shown). Figure 3 shows that few videos included more than three different spatial scale levels (Fig. 3a), and it was uncommon to convey transitions between scale levels (Fig. 3c). A large fraction of videos lacked any indication of temporal scale (Fig. 3b).

3.3 Misconceptions

All included misconceptions were conveyed less frequently than the key concepts, except for the key concepts *number of traits* and *negative mutations*. In fact, most videos (almost 80%) did not communicate any misconceptions, and it was uncommon that more than one was conveyed (Fig. 3d). The most common incorrect explanations were that variation occurs in response to *need* and that *only beneficial traits are inherited* (Fig. 2).

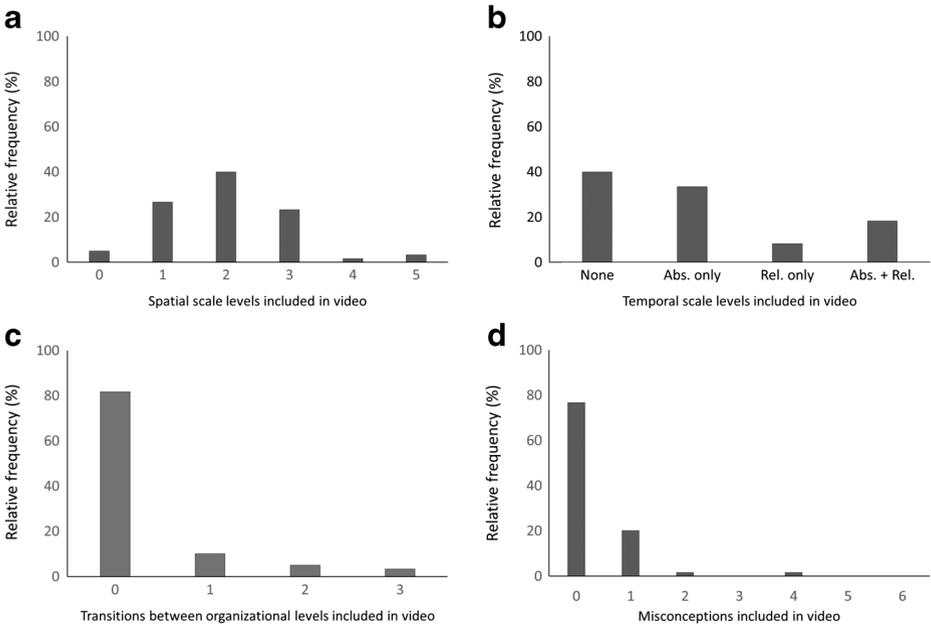


Fig. 3 Relative frequency (%) distribution for the number of variables that were coded as present for each of the categories **a** spatial-scale levels, **b** temporal-scale levels, **c** transitions between organizational levels, and **d** misconceptions. In **b**, the original variables have been collapsed into two variables: relative time (occurrence of the *generations* variable) and absolute time (occurrence of any other temporal scale levels variable)

3.4 Organismal Context

Notably, most (70%) of the videos concerned evolution in animals other than humans, while 25% included human evolution. Symbolic organisms were used in 10% and bacteria in 18% of the videos, while evolution in plants was only represented in 8% of them.

3.5 Relations Across Categories

The clustering of 22 variables yielded a four-cluster solution. The clusters are displayed in Table 3.

The video-based clustering yielded four clusters, containing 35, 20, 4, and 1 video(s) each. Table 4 characterizes the clusters in terms of the variables that appeared in at least 50% of the included videos. The majority of the expressed misconceptions are found in cluster 1 (13 of 18 occurrences in total), whereas none are found in cluster 3 or 4. Note that misconceptions do not appear in the table since they were relatively rare as individual variables (relative presence of 0–10%), although 20% of the total sample contained at least one misconception (Fig. 3d).

Table 5 characterizes the four clusters from the video-based clustering based on length and stated audience. Clusters 1 and 2 exhibited similar minimum, average, and maximum video lengths. It can be seen that no video shorter than 4 min was included in clusters 3 and 4. Two videos stated the intended audience as “educators, students and parents”. One video was intended for teacher professional development and 12 were directed toward students (5 to high school, 1 to middle school, and 6 unspecified). Most videos (45 out of 60) did not state any intended specific audience or general target group.

4 Discussion

A large body of research has identified several crucial aspects of natural selection that are vital for understanding the central theory of life (Gregory 2009; Smith 2010b). However, little attention has been paid to the expression of these aspects in online videos purporting to explain natural selection to the public and/or students. Therefore, this study explored the extent to which selected conceptual aspects important for natural selection were expressed in a sample of videos available online. The results showed that some variables were presented much more consistently than others across the videos. For example, the spatial and temporal scales that

Table 3 Results from the variable-based clustering

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Key concepts	Individual variation Inherited variation Differential fitness Limited survival Change in population Speciation	Origin of variation	Limited resources Extinction	
Level(s) of organization (space)	Individual Population	Molecule Genome Cell to individual	Species	
Level(s) of organization (time)	Generations	Years		
Randomness/probability	Probability	Randomness		
Organismal context		Humans	Animals	Bacteria

Table 4 Results from the video-based clustering. Variables that appear in 50% or more of the included videos are shown

	Cluster 1 (n = 35)	Cluster 2 (n = 20)	Cluster 3 (n = 4)	Cluster 4 (n = 1)
Key concepts		Individual variation Inherited variation Differential fitness Limited survival Change in population	Origin of variation Individual variation Inherited variation Differential fitness Limited survival Change in population	Origin of variation Individual variation Change in population
Level(s) of organization (space)	Individual	Individual Population	Speciation Molecule Genome Cell to individual Individual Population Species	Molecule Genome Gene to protein Cell to individual Individual
Level(s) of organization (time)		Years Generations	Years Generations	Days Years
Randomness/probability		Probability	Randomness	Randomness Probability
Organismal context	Animals	Animals	Animals Humans	Bacteria

were most often portrayed concern the size and generation time of meter-scale organisms, which are organizational levels that humans can directly relate to. However, variation originates through mutations in our cells that occur randomly at much shorter temporal scales and smaller spatial scales. These processes are expressed in the videos less often, and their connections with macroscopic and directly observable phenomena are even rarer.

The present study explored the type of content that is conveyed to viewers that search for accessible information about natural selection online. However, when using videos and animations in teaching, one should acknowledge the importance of scaffolding, given that watching videos is generally a passive endeavor. Chi and Wylie (2014) have provided valuable advice on how students can work more actively with videos based on their ICAP (Interactive Constructive Active Passive) framework. Their approaches include manipulation of videos,

Table 5 Length and stated audiences for the videos in the respective clusters

	Cluster 1 (n = 35)	Cluster 2 (n = 20)	Cluster 3 (n = 4)	Cluster 4 (n = 1)	Total
Average length (min/max) minutes:seconds	4:37 (0:45/17:21)	4:49 (0:30/17:39)	7:36 (4:22/11:48)	7:48	4:56 (0:30/17:39)
Audience					
Unspecified	30	12	2	1	45
Educators, students, parents	0	2	0	0	2
Teachers	1	0	0	0	1
Students in general	3	1	2	0	6
High school	1	4	0	0	5
Middle school	0	1	0	0	1

actively comparing the contents to other learning material, and discussing and debating with peers (Chi and Wylie 2014).

An important note is that videos showing a plurality of variables are not necessarily more suitable for learning or teaching. We assessed the extent to which the variables were present in the sampled videos but do not make any claims with regard to how well they are conveyed. In addition, the instructional usefulness of different videos is likely to differ depending on the previous knowledge and educational level of the viewer. For example, a middle school pupil would probably not require the necessarily complex picture that would arise from integrating all included variables in a single video. A more advanced learner might possibly also prefer a set of shorter videos that focus on different concepts. However, these remain hypothetical questions that could be investigated in future studies. Nevertheless, irrespective of a viewer's pre-knowledge, it would be very difficult to learn a specific concept from a video if it is not present at all.

4.1 Key Concepts

A striking finding is that a fifth (12 of 60) of the included videos purporting to explain natural selection did not include even one of the key concepts used in the study. This finding alone calls for caution when searching for educational material online. Moreover, the largest cluster of videos, comprising 35 videos, was characterized by a low presence of the key concepts, implying that it could be challenging to acquire a full understanding of natural selection from these videos.

The most frequently appearing key concepts (*individual variation, inherited variation, differential fitness, limited survival, and change in population*) co-occur in the variable-based clustering. Together, these key concepts could convey the message that there is natural variation within a population, that the variation is inherited by the next generation and that some individuals are fitter (i.e. more likely to survive and reproduce) than others, which essentially serves as a basic description of natural selection. A more complete account would require the key concepts to be complemented with the variation-generating process (*origin of variation, mutation*) and that changes in populations sometimes lead to the development of new species and extinction of others. However, the origin of variation is conveyed to a smaller extent and is found in a different cluster of variables than the abovementioned key concepts. The origin of variation variable clusters with, for example, the variables randomness and genomic or molecular scale. Together, this reflects the subcellular nature of the processes that lead to new variation. All of the above key concepts appear together in the four videos (7% of all videos) making up one of the video clusters.

Few videos (10%) illustrate *more than one trait* and even fewer illustrate *negative mutations* (7%). These two concepts are not generally considered key concepts of natural selection. However, only showing one trait at a time might give the impression that only one trait varies at a time, an idea that may lead to several misconceptions (e.g. Redfield 2012). Ignoring negative mutations could reinforce the misconceptions held by some learners that mutations are always beneficial and not random (Gregory 2009). The low frequencies of multiple traits and negative mutations may simply reflect a desire to reduce complexity in the videos, but whether this makes the narratives easier to follow or induces misconceptions should be considered by anyone contemplating using them for educational purposes.

Given the postulated importance of the principles (*variation, inheritance and selection*) for understanding natural selection (e.g. Tibell and Harms 2017), key concepts belonging to all three of them were conveyed in surprisingly few of the videos. While 32% included all principles according to the relaxed approach for regarding a principle as present, a mere 2% included them under the strict approach where all underlying key concepts are required

(Table 2). This implies that a learner would probably need to view several of these videos to be exposed to all three principles.

4.2 Threshold Concepts

The most frequently represented *levels of organization* in space and time in the videos are individuals/populations and years/generations, respectively (note that the reliability fell below the agreed cutoff level for events at the individual level (.4) and slightly below for generations (.64)). Furthermore, time is more commonly expressed in years than in generations. This is somewhat unfortunate since time in terms of generations, as a relative time-scale, provides much richer information for the learner and facilitates comparison of organisms with different generation times (e.g. Catley and Novick 2009). A suggested method to overcome the vast differences in space and time scales is to study evolution through an example organism with short generation time, such as a bacterium (e.g. Bohlin and Höst 2015).

The processes underlying the expression of traits are sometimes presented in the videos, but the mechanisms whereby genes are decoded, resulting in formation of proteins that participate in the cellular functions that ultimately lead to individual characteristics, are very seldom explained. The low frequency of videos including small (relative to human perception) spatial and temporal scales is potentially problematic because it does not provide learners with an appreciation of the fundamental similarities between different groups of organisms that arise from their common ancestry. These similarities include, for example, the roles of molecules such as DNA, RNA, and proteins, and of processes such as replication. These common life features and processes are vital components of evolutionary processes, and their ubiquity contributes to the high explanatory value of evolution in biology. Previous studies have revealed differences in the extent that experts and novices apply theories to a range of evolutionary problems (Nehm and Ridgway 2011). This is not surprising given the problem of transfer in learning situations (Day and Goldstone 2012). Thus, treating spatial scales as a threshold concept offers a tempting way to convey the similarity of apparently dissimilar living organisms and overcome part of the transfer problem. Furthermore, videos enabling comparison and appreciation of these similarities could possibly play important roles in successful instructional strategies for teaching natural selection.

The only key concept that clustered together with threshold concepts in the variable-based clustering was the origin of variation, which co-occurred together with randomness and smaller spatial scales. This could be explained under the premise that mutations are the ultimate source of novel variation within a population—the mutations are random with respect to the environment the organisms inhabit. In addition, many mutations occur simply due to errors during DNA replication. Thus, randomness accompanies many crucial elements of evolution, but as shown by Gregory (2009), many studies indicate that misunderstandings of evolution are frequently linked to confusion regarding the role of randomness. For example, students tend to believe either that evolution is random or that it is driven by teleological principles (Bizzo 1994). Randomness is conveyed less frequently than probability, a concept that is highly relevant for discussing how selection acts on existing variation (Tibell and Harms 2017). This may explain that probability is present to a higher degree than randomness in video cluster 2, wherein videos typically contain most key concepts except origin of variation. To our knowledge, no published studies have explored whether the observed underrepresentation of random variation-generating processes in the studied evolution videos is mirrored in other types of learning material. However, according to Garvin-Doxas and Klymkowsky (2008), students have

difficulties in understanding the random components of evolution, partly due to the preconception that random processes are inefficient. Accordingly, Robson and Burns (2011) have found that interventions targeting the concept of randomness with regard to mutations in evolution have positive effects on student knowledge. Thus, the concept of randomness and its relation to natural selection warrants more attention by designers of learning material intended to enhance understanding of natural selection and evolution.

4.3 Misconceptions

The notion that evolution is driven by *need* is a misconception associated with an erroneous understanding of the random processes of evolution. It is also the most commonly expressed misconception in the examined videos (about 10% actually supported this view). While need is sometimes used as a metaphor for fitness, i.e. factors relevant to survival, it can be problematic for novices who lack a sound theoretical understanding of natural selection and thus interpret the metaphor literally. Failure to appreciate the metaphorical use of need can potentially lead novices to believe that a need for change in a particular trait is enough to cause the appearance of novel variation that satisfies the needs of an individual, a population or even a species. The second most common misconception that might be reinforced in the analyzed videos is that *only beneficial traits are inherited*. A way to overcome this problem might be to include examples of mutations leading to reductions in evolutionary fitness in the videos. It was not surprising that the majority of the observed misconceptions were contained in the first video cluster, which was characterized by a low frequency of key concepts. The relatively large size of the cluster (35 out of 60 videos) indicates that educators should examine videos with caution before integrating them as teaching resources. In this regard, explicitly addressing common misconceptions has been shown to be a successful teaching strategy (e.g. Nehm and Schonfeld 2007). However, this would require that misconceptions and correct ideas are addressed simultaneously, which was clearly not the case in the present sample of videos. Therefore, exploring if and how videos teach about, rather than perpetuate, misconceptions presents an interesting approach for future studies.

4.4 Organismal Context

The skewed distribution of types of organisms represented in videos is interesting in light of earlier findings that learners' explanations of biological change (Southerland et al. 2001) and evolution (Nehm and Ha 2011; Opfer et al. 2012) differ depending on the context and type of organism. Thus, as the explanatory coherence of natural selection across different contexts is reportedly weak among learners, the predominant use of animals in videos is a potential concern. Even though an animal context could be beneficial with regard to familiarity, for example, the lack of alternative contexts might present problems for understanding how natural selection transcends organismal contexts. The possibility that diverse examples with regard to organismal context and trait gain/loss could help learners see beyond surface features is therefore a hypothesis that deserves more attention. This is supported by findings that students perform worse on test items where plants have been used as models than on items based on animal models (Ha et al. 2006). In this regard, we note that previous studies on conceptual understanding of evolution have focused mostly on animals, which might indicate that the observed bias is also present among science education researchers. The only types of organisms

explicitly considered in this study are humans, animals, plants, bacteria and symbolic organisms, although others (e.g. fungi or single-celled eukaryotes) could contribute to a more complete picture. However, we did not observe any representations of other types of organisms in our set of videos. The finding that the variable bacteria as organismal context did not cluster with other variables in the variable-based clustering is intriguing and indicates that the use of bacteria in videos does not predict what other variables are present or absent in the videos. This finding calls for further research into the potential of using a bacterial context when teaching natural selection.

4.5 Audience and Length

Given the similar values across clusters, the length of the videos does not seem to be an indicator of what types of variables were conveyed. The average length of the videos sampled was 4:56 min, which resembles the average length (4.4 min) of online content videos (Lella 2014). Welbourne and Grant (2016) found no correlation between video length and popularity of scientific videos and conclude that content creators do not have reasons to assume any length to be more appropriate than another. Our findings seem to extend this conclusion to educators looking for useful videos to include in their teaching. Regarding the intended audiences, it is a striking finding that 45 out of 60 videos do not specify the level of audience at all. Among those that do, only six videos specify the audience in terms of high or middle school.

4.6 Limitations

When considering time scales, the variable with the longest reach in absolute time was measured in years. This made it impossible to differentiate between videos that contained information concerned with a few years to those involving deep time during the analysis. This is something that should be considered when planning future studies.

Our sampling procedure intended to mimic frequently used search strategies and mined videos in a realistic way as experienced by typical users. Moreover, there were certain sources (such as ed.ted.com) that were deliberately included in the sample due to their popularity and perceived trustworthiness. These methodological choices restrict the reproducibility of the sampling, and therefore, the results cannot be generalized onto what aspects of natural selection that are conveyed by all existing educational videos. However, searching for and watching videos is largely a metacognitive activity and videos that are recommended by a website's internal system, rather than those appearing as a direct result from the used search terms, is a strong factor in determining what videos are viewed by persons searching for biology-related content (Lei et al. 2015). Therefore, although there are clear limitations associated with our sampling procedure, our results might provide a more accurate depiction of what a user might encounter.

Evolution consists of a number of linked mechanisms (such as genetic drift, gene flow, and natural selection) both at the macro and micro-level. Since the criteria utilized in this study were chosen with regard to natural selection, we cannot draw conclusions about how the included videos conveyed information on other evolutionary mechanisms. However, threshold concepts such as randomness and spatial scale have profound implications for understanding both genetic drift and gene flow (e.g. Price et al. 2014), indicating that this is an intriguing future research direction.

Although this study provides indications of the extent to which important aspects of evolution are presented in online videos, no inferences can be made regarding the quality of the presentations. It should be noted that there is no guarantee that students will understand a particular concept simply through viewing an animation that portrays it. While evaluating the effectiveness of different design choices presents an interesting research opportunity, it is beyond the scope of this study.

4.7 Future Directions

The present study represents a first step in the characterization of online videos purporting to explain natural selection. Given its exploratory nature, several future research opportunities have emerged. Follow-up studies could further develop strategic sampling methods applicable to modern streaming sites and employ more standardized search criteria to increase the generalizability of the findings. Furthermore, there are many aspects pertaining to the educational value of videos that deserve investigation—How could threshold concepts be efficiently visualized and related to key concepts for different audiences? What affordances do different organismal contexts have on learners' perceptions of natural selection, and how could transfer across taxa be facilitated through videos and animations? Lastly, what kinds of scaffolding are of an educational advantage when using videos as a teaching tool for different age groups?

4.8 Conclusions

We have provided a first record of the concepts related to natural selection that are conveyed by online videos purporting to explain the theory of evolution. Our results indicate that random factors, especially in the generation of variation, are underrepresented in relation to non-random processes affecting individual or population level phenomena. Given earlier findings that conceptualizing random factors is important for evolutionary understanding, this has profound implications for teaching and designing videos. Thus, more research on the nature of the relation between randomness and understanding of natural selection is urgently warranted. Furthermore, a significant proportion of the analyzed videos do not fully address the basic principles of variation, inheritance, and selection. It is possible that not all videos were originally intended to convey the basic mechanism of evolution by mutation and natural selection or that the designers have assumed that learners already understand some of these basic principles. Nevertheless, it is remarkable that few of the videos exploit opportunities of the media to communicate basic evolutionary principles and display the complexity of natural selection. Lastly, the predominant use of animals as organismal context is striking, and the possibility that this bias may impede learners' perception of evolution as a universal explanation for phenomena across all of biology is a question that should be pursued in the future.

Acknowledgments The authors acknowledge the EvoVis research group for many and profound discussions. Three anonymous reviewers provided insightful comments and valuable suggestions. We are also very grateful to Dr. John Blackwell and Dr. Konrad Schönborn for language review and valuable comments of the paper.

Funding Information The Swedish Research Council (VR 2012:5344, LT) supported this research.

Appendix

Table 6 Complete list of included variables, their operationalization, relative frequency of occurrence in the videos, and reliability (Krippendorff's alpha coefficients)

Variable	Operationalization	Relative frequency of occurrence (%)	Reliability	Agreement (%)
Variation				
1. Origin of variation	Does the video in any way depict that variation originates at the genetic level? For example, through mutations, genetic recombination or horizontal gene transfer.	31.7	.92	97%
2. Individual variation	Does the video in any way depict that individuals in a population are not identical but vary in different characteristics?	48.3	.81	91%
3. Differential fitness	Does the video show that individuals have different chances of survival and reproduction in their environment, due to their varied characteristics?	58.3	.81	91%
Inheritance				
4. Inherited variation	Does the video depict that the traits/characteristics are inherited from parents to progeny?	36.7	.88	94%
Selection				
5. Limited resources	Does the video show that natural resources (e.g. space and nutrients) are limited?	16.7	.92	97%
6. Limited survival	Does the video show that not all individuals of a generation survive to reproduce (e.g. due to limited resources or predator attacks)?	41.7	.94	97%
7. Change in population	Does the video show that individuals with favorable characteristics gradually come to dominate populations over generations?	41.7	.82	91%
8. Speciation	Does the video show that separated populations may diverge sufficiently to become separate species?	16.7	.84	97%
Additional variables				
9. Number of traits	Does the video follow the evolution of more than one trait?	10.0	.72	94%
10. Negative mutations	Does the video show the occurrence of mutations leading to reductions in individuals' evolutionary fitness?	6.7	1.00	100%
11. Extinction	Does the video show that all individuals of a species may die off? (e.g. most species of dinosaurs).	13.3	.72	94%
Spatial scale				
12. Molecule	Does the video show events at the molecular scale (1–100 nm), e.g. changes at the nucleotide level (mutations)?	13.3	.72	94%
13. Genome	Does the video show events at the genome scale (100 nm–10 μ m), e.g. changes in chromosomes and genes?	30.0	.72	91%
14. Individual	Does the video show events at the scale of an individual organism (10 μ m–50 m)?	75.0	.40	76%

Table 6 (continued)

Variable	Operationalization	Relative frequency of occurrence (%)	Reliability	Agreement (%)
15. Population	Does the video show events at the level of populations (all organisms of a given group or species in a given geographical area capable of interbreeding)?	43.3	.82	91%
16. Species	Does the video show events at species level (e.g. growing cladograms)?	38.3	.68	88%
Transition between organizational levels				
17. Gene-protein	Does the video show that mutations lead to changes in protein structure and function?	8.3	0	97%
18. Protein-cellular function	Does the video show that changes in protein structure and function are directly coupled to changes in cellular function?	3.3	N/A	100%
19. Cellular function-individual characteristics	Does the video show the coupling between changes at the cellular level and characteristics of the individual?	18.3	.72	94%
Temporal scale				
20. Shorter than seconds	Does the video provide information about temporal changes at time scales shorter than seconds (e.g. microseconds)?	1.7	1.00	100%
21. Seconds	Does the video provide information about temporal changes in seconds?	0.0	N/A	100%
22. Minutes/hours	Does the video provide information about temporal changes in minutes or hours?	0.0	N/A	100%
23. Days	Does the video provide information about temporal changes in days?	1.7	0	97%
24. Years	Does the video provide information about temporal changes in years?	50.0	.72	88%
25. Generations	Does the video provide information about temporal changes indicated in numbers of generations?	26.7	.64	88%
Randomness				
26. Randomness	Does the video convey the molecular level random factors involved in the generation of variation?	23.3	.91	97%
Probability				
27. Probability	Does the video indicate factors affecting individuals' probabilities of survival and reproduction?	31.7	.60	82%
Misconceptions				
28. Need	Does the variation appear in response to a change in the environment?	10.0	.79	97%
29. Anthropolomorphism	Do entities behave as if they had human intentions and ability to plan ahead?	3.3	.66	97%
30. Only good traits are inherited	Do all individuals that do not have illustrated beneficial trait(s) die?	8.3	.53	91%
31. Natural selection as an event	Do major changes occur instantaneously, in one generation, not through gradual shifts in genotype and phenotype frequencies over time?	3.3	1.00	100%

Table 6 (continued)

Variable	Operationalization	Relative frequency of occurrence (%)	Reliability	Agreement (%)
32. Acquired traits are inherited	Are traits that agents in the video develop during their lifetime passed on to their offspring?	0	N/A	100%
33. Essentialism	Are changes in species displayed as simultaneous transformations of the essential nature of all the individuals?	5.0	1.00	100%
Organismal context				
34. Bacteria	Does the video show evolution of bacteria?	18.3	1.00	100%
35. Animals	Does the video show evolution of animals, excluding humans?	70.0	.73	88%
36. Humans	Does the video show evolution of humans?	25.0	.82	94%
37. Plants	Does the video show evolution of plants (including flowers, trees, algae etc.)?	8.3	1.00	100%
38. Symbolic organisms	Does the video show evolution of a symbolic organism?	10.0	.79	97%

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Aldridge, M., & Dingwall, R. (2003). Teleology on television? *European Journal of Communication*, 18(4), 435–453.
- Anderberg, M. R. (2014). *Cluster analysis for applications*. New York: Academic Press, Inc..
- Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the conceptual inventory of natural selection. *Journal of Research in Science Teaching*, 39(10), 952–978.
- Andrews, T. M., Price, R. M., Mead, L. S., McElhinny, T. L., Thanukos, A., Perez, K. E., et al. (2012). Biology undergraduates' misconceptions about genetic drift. *CBE—Life Sciences Education*, 11(3), 248–259.
- Bishop, B. A., & Anderson, C. W. (1990). Student conceptions of natural selection and its role in evolution. *Journal of Research in Science Teaching*, 27(5), 415–427.
- Bizzo, N. M. V. (1994). From down house landlord to Brazilian high-school-students—what has happened to evolutionary knowledge on the way. *Journal of Research in Science Teaching*, 31(5), 537–556.
- Bohlin, G., & Höst, G. H. (2015). Evolutionary explanations for antibiotic resistance in daily press, online websites and biology textbooks in Sweden. *International Journal of Science Education, Part B: Communication and Public Engagement*, 5(4), 319–338.
- Brown, J. H., Stevens, G. C., & Kaufman, D. M. (1996). The geographic range: size, shape, boundaries and internal structure. *Annual Review of Ecology and Systematics*, 27, 597.
- Catley, K. M., & Novick, L. R. (2009). Digging deep: exploring college students' knowledge of macroevolutionary time. *Journal of Research in Science Teaching*, 46(3), 311–322.
- Cheek, K. A. (2010). Why is geologic time troublesome knowledge? In J. H. F. Meyer, R. Land, & C. Baillie (Eds.), *Threshold concepts and transformational learning* (pp. 117–129). Rotterdam: Sense Publishers.
- Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243.
- Davies, P. (2006). Threshold concepts: how can we recognize them. In J. H. F. Meyer & R. Land (Eds.), *Overcoming barriers to student understanding: threshold concepts and troublesome knowledge* (pp. 70–84). Abingdon: Routledge.
- Day, S. B., & Goldstone, R. L. (2012). The import of knowledge export: connecting findings and theories of transfer of learning. *Educational Psychologist*, 47(3), 153–176.
- Depew, D. (2013). Conceptual change and the rhetoric of evolutionary theory: 'force talk' as a case study and challenge for science pedagogy. In K. Kampourakis (Ed.), *The philosophy of biology: a companion for educators* (pp. 121–144). Springer Netherlands.
- Dingwall, R., & Aldridge, M. (2006). Television wildlife programming as a source of popular scientific information: a case study of evolution. *Public Understanding of Science*, 15(2), 131–152.
- Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher*, 35(3), 125–129.
- Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). Chichester: John Wiley & Sons.
- Fletcher, J. D., & Tobias, S. (2005). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (Vol. 1, pp. 117–133). Cambridge: Cambridge University Press.
- Futuyama, D. J. (1995). The uses of evolutionary biology. *Science*, 267(5194), 41–42.
- Garvin-Doxas, K., & Klymkowsky, M. W. (2008). Understanding randomness and its impact on student learning: lessons learned from building the biology concept inventory (BCI). *CBE—Life Sciences Education*, 7, 227–233.
- Gilbert, J. K. (Ed.). (2005). *Visualization in science education*. Dordrecht: Springer.
- Godfrey-Smith, P. (2007). Conditions for evolution by natural selection. *The Journal of Philosophy*, 104(10), 489–516.
- Gregory, T. R. (2009). Understanding natural selection: essential concepts and common misconceptions. *Evolution: Education and Outreach*, 2(2), 156–175.
- Ha, M.-S., Lee, J.-K., & Cha, H.-Y. (2006). A cross-sectional study of students' conceptions on evolution and characteristics of concept formation about it in terms of the subjects: human, animals and plants. *Journal of The Korean Association for Science Education*, 26(7), 813–825.
- Hamid, J. S., Meaney, C., Crowcroft, N. S., Granerod, J., & Beyene, J. (2010). Cluster analysis for identifying sub-groups and selecting potential discriminatory variables in human encephalitis. *BMC Infectious Diseases*, 10(1), 364.

- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*(1), 77–89.
- Hiatt, A., Davis, G. K., Trujillo, C., Terry, M., French, D. P., Price, R. M., & Perez, K. E. (2013). Getting to evolve: concepts and challenges for students learning evolutionary developmental biology. *CBE—Life Sciences Education, 12*(3), 494–508.
- Islam, M. M., Valderas, J. M., Yen, L., Dawda, P., Jowsey, T., & McRae, I. S. (2014). Multimorbidity and comorbidity of chronic diseases among the senior Australians: prevalence and patterns. *PLoS One, 9*(1), e83783.
- Johnstone, A. H. (1982). Macro- and microchemistry. *School Science Review, 64*, 377–379.
- Jördens, J., Asshoff, R., Kullmann, H., & Hammann, M. (2016). Providing vertical coherence in explanations and promoting reasoning across levels of biological organization when teaching evolution. *International Journal of Science Education, 38*(6), 960–992.
- Kampourakis, K. (2015a). Distorting the history of evolutionary thought in conceptual development research. *Cognitive Science, 39*(4), 833–837.
- Kampourakis, K. (2015b). The need for interdisciplinary dialog in evolution education: a comment on the responses by Ware & Gelman and Shtulman. *Cognitive Science, 39*(4), 846–848.
- Krippendorff, K. (2004). *Content analysis: an introduction to its methodology* (2nd ed.). Thousand Oaks: Sage.
- Lee, S. W. Y., & Tsai, C. C. (2013). Technology-supported learning in secondary and undergraduate biological education: observations from literature review. *Journal of Science Education and Technology, 22*(2), 226–233.
- Lei, P. L., Sun, C. T., Lin, S. S., & Huang, T. K. (2015). Effect of metacognitive strategies and verbal-imagery cognitive style on biology-based video search and learning performance. *Computers & Education, 87*, 326–339.
- Lella, A. (2014). comScore releases January 2014 U.S. online video rankings. Retrieved from <http://www.comscore.com/Insights/Press-Releases/2014/2/comScore-Releases-January-2014-US-Online-Video-Rankings>
- Leonard, M. J., Kalinowski, S. T., & Andrews, T. C. (2014). Misconceptions yesterday, today, and tomorrow. *CBE—Life Sciences Education, 13*, 179–186.
- Lewis, J., & Kattmann, U. (2004). Traits, genes, particles and information: re-visiting student's understandings of genetics. *International Journal of Science Education, 26*(2), 195–206.
- Lewontin, R. C. (1970). The units of selection. *Annual Review of Ecology and Systematics, 1*, 1–18.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayr, E. (1982). *The growth of biological thought: diversity, evolution and inheritance*. Cambridge: Harvard University Press.
- Mayr, E. (2001). *What evolution is. Basic books*. New York: Perseus Book Group.
- Meyer, J. H. F., & Land, R. (2003). Threshold concepts and troublesome knowledge: linkages to ways of thinking and practising within the disciplines. In *Enhancing teaching-learning environments in undergraduate courses* Occasional report 4.
- Meyer, J. H. F., & Land, R. (2005). Threshold concepts and troublesome knowledge (2): epistemological considerations and a conceptual framework for teaching and learning. *Higher Education, 49*(3), 373–388.
- Milne, M. J., & Adler, R. W. (1999). Exploring the reliability of social and environmental disclosures content analysis. *Accounting, Auditing & Accountability Journal, 12*(2), 237–256.
- Morris, D. W., & Lundberg, P. (2011). *Pillars of evolution: fundamental principles of the eco-evolutionary process*. Oxford: Oxford University Press.
- Nehm, R. H., & Ha, M. (2011). Item feature effects in evolution assessment. *Journal of Research in Science Teaching, 48*(3), 237–256.
- Nehm, R. H., Rector, M. A., & Ha, M. (2010). “Force-talk” in evolutionary explanation: metaphors and misconceptions. *Evolution: Education and Outreach, 3*(4), 605–613.
- Nehm, R. H., & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *Bioscience, 57*(3), 263–272.
- Nehm, R. H., & Ridgway, J. (2011). What do experts and novices “see” in evolutionary problems? *Evolution: Education and Outreach, 4*(4), 666–679.
- Nehm, R. H., & Schonfeld, I. S. (2007). Does increasing biology teacher knowledge of evolution and the nature of science lead to greater preference for the teaching of evolution in schools? *Journal of Science Teacher Education, 18*, 699–723.
- Nehm, R. H., & Schonfeld, I. S. (2008). Measuring knowledge of natural selection: a comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching, 45*(10), 1131–1160.
- Nieswandt, M., & Bellomo, K. (2009). Written extended-response questions as classroom assessment tools for meaningful understanding of evolutionary theory. *Journal of Research in Science Teaching, 46*(3), 333–356.
- Olson, S. (2012). Learning about evolution: the evidence base. In J. B. Labov (Ed.), *Thinking evolutionarily: evolution education across the life sciences: summary of a convocation* (pp. 25–31). Washington, D.C.: The National Academies Press.
- Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: knowing what students know about evolution. *Journal of Research in Science Teaching, 49*(6), 744–777.

- Ortiz-Cordova, A., Yang, Y., & Jansen, B. J. (2015). External to internal search: associating searching on search engines with searching on sites. *Information Processing & Management*, 51(5), 718–736.
- Pallant, A., & Tinker, R. F. (2004). Reasoning with atomic-scale molecular dynamic models. *Journal of Science Education and Technology*, 13, 51–66.
- Phillips, L. M., Norris, S. P., & Macnab, J. S. (2010). *Visualization in mathematics, reading and science education* (Vol. Vol. 5). Dordrecht: Springer Netherlands Retrieved from <http://link.springer.com/10.1007/978-90-481-8816-1>.
- Price, R. M., Andrews, T. C., McElhinny, T. L., Mead, L. S., Abraham, J. K., Thanukos, A., & Perez, K. E. (2014). The genetic drift inventory: a tool for measuring what advanced undergraduates have mastered about genetic drift. *CBE—Life Sciences Education*, 13(1), 65–75.
- Redfield, R. J. (2012). Why do we have to learn this stuff?—a new genetics for 21st century students. *PLoS Biology*, 10(7), e1001356.
- Robson, R. L., & Burns, S. (2011). Gain in student understanding of the role of random variation in evolution following teaching intervention based on Luria-Delbruck experiment. *Journal of Microbiology & Biology Education*, 12(1), 3–7.
- Ross, P. M., Taylor, C. E., Hughes, C., Whitaker, N., Lutze-Mann, L., Kofod, M., & Tzioumis, V. (2010). Threshold concepts in learning biology and evolution. *Biology International*, 47, 47–54.
- Rundgren, C.-J., & Tibell, L. A. E. (2010). Critical features of visualizations of transport through the cell membrane: an empirical study of upper secondary and tertiary students' meaning-making of a still image and an animation. *International Journal of Science and Mathematics Education*, 8(2), 223–246.
- Shtulman, A. (2006). Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology*, 52, 170–194.
- Shtulman, A., & Schulz, L. (2008). The relation between essentialist beliefs and evolutionary reasoning. *Cognitive Science*, 32(6), 1049–1062.
- Sinatra, G. M., Brem, S. K., & Evans, E. M. (2008). Changing minds? Implications of conceptual change for teaching and learning about biological evolution. *Evolution: Education and Outreach*, 1(2), 189–195.
- Smith, M. U. (2010a). Current status of research in teaching and learning evolution: I. Philosophical/epistemological issues. *Science & Education*, 19, 523–538.
- Smith, M. U. (2010b). Current status of research in teaching and learning evolution: II. Pedagogical issues. *Science & Education*, (19), 539–571.
- Southerland, S. A., Abrams, E., Cummins, C. L., & Anzelmo, J. (2001). Understanding students' explanations of biological phenomena: conceptual frameworks or p-prims? *Science Education*, 85(4), 328–348.
- Spiegel, A., Evans, M., Gram, W., & Diamond, J. (2006). Museum visitors' understanding of evolution. *Museums & Social Issues*, 1(1), 69–86.
- Taylor, C. (2006). Threshold concepts in biology: do they fit the definition? In J. H. F. Meyer & R. Land (Eds.), *Overcoming barriers to student understanding: threshold concepts and troublesome knowledge* (pp. 87–99). Abingdon: Routledge.
- Tibell, L.A.E., & Harms, U. (2017). Biological principles and threshold concepts for understanding natural selection—implications for developing visualizations as a pedagogic tool. *Science & Education*, <https://doi.org/10.1007/s11191-017-9935-x>.
- Understanding Evolution. (2017). University of California Museum of Paleontology. 26 January 2017 <<http://evolution.berkeley.edu/>>.
- van Dijk, E. M., & Reydon, T. A. C. (2010). A conceptual analysis of evolutionary theory for teacher education. *Science & Education*, 19, 655–677.
- Welbourne, D. J., & Grant, W. J. (2016). Science communication on YouTube: factors that affect channel and video popularity. *Public Understanding of Science*, 25(6), 706–718.