Live-action Communication Design: A Technical How-To Video Case Study

Per Erik Eriksson and Yvonne Eriksson

School of Humanities and Media Studies, Dalarna University; School of Innovation, Design and Engineering, Mälardalen University

ABSTRACT
This case study is based on a research through design project (RTD) that focuses on a technical communication video of the live-action format. It investigates the usability and design-implications of a live-action how-to video, by means of analyzing user-centered data such as YouTube analytics data, usability, and comprehension assessments. In the study, four key live-action video affordances are identified: verifiability, comparability, recordability, and visibility. The identification of these affordances when related to the users’ assessments resulted in several design implementations that would warrant sought-for communication efficacies. Findings show that some assumed efficacies appear to be mitigated by the complexity and the density of the video information. One implication of this is that the implementation of conventional video editing techniques and the addition of on-screen text that serve to make content briefer and more concise into instructional live-action videos requires the technical communicator’s careful consideration.

KEYWORDS
Assessment; design of communication; digital technologies; knowledge management; research methods; usability studies; visual rhetoric/visualization techniques

Introduction
This article is engaging matters of professional practice by looking at one technical communication (TC) video, its associated videography-driven design process and the way the video supports its usership in a real-world assembly and construction scenario. Technical communication videos are forms of information and knowledge support tools. As such, according to William B. Rouse (2007), they should be “useful,” to the extent to which information and knowledge help users to pursue their intentions. They should also be “usable,” to the extent to which information and knowledge are easily accessed, digested and applied, and they should be “urgent,” to the extent to which information and knowledge help users to pursue near-term plans (p. 222). In a similar vein to Rouse and other human-centered design (HCD) researchers, the research approach that underpins the case study presented in this article considers users’ evaluations of a critical component of prototype development. This case study is also the result of a research through design project (RTD). This means that the authors of this article developed the video prototype that is the focal point of this case study and that this video prototype and its associated perceptual affordances (cf. Teston, 2012) play a central role in the research and knowledge generating process.

How-to videos on YouTube
In this moving, image-crazed world, TC videos have become ubiquitous and surprisingly diverse. However, we suggest that assembly videos, tutorial videos, standard operating procedural videos,
reference manual videos, and other forms of technical documentation videos have a number of things in common. First, we note that most of them may be defined as “how-to” videos and commonly are among the 585 million (!) instructional videos that can be accessed on YouTube on typing the keywords how to (consulted in March 2018). Second, with the knowledge that (computer-generated) animations are growing in popularity, we note that many of these videos include live-action video content one way or the other. In other words, to various degrees many how-to videos include recorded video that displays a “live” instruction that unfolds in front of a smartphone or a video camera.

What becomes evident when using instructional live-action videos accessible via YouTube is that they represent an extremely uneven quality or, more to the point, uneven quality definitions. In spite of a few concerted research efforts into the communication efficacies of the live-action format (Arguel & Jamet, 2009; Boucheix & Forestier, 2017; Chirumalla, Eriksson, & Eriksson, 2015; Ganier & de Vries, 2016; Marcus, Cleary, Wong, & Ayres, 2013; van Hooijdonk & Krahmer, 2008; Watson, Butterfield, Curran, & Craig, 2010), and though many redeeming communication aspects of good TC videos are the same as in effective written procedures (Alexander, 2013; Morain & Swarts, 2012), the instructional live-action format presents unique communication challenges that remain to be satisfactorily understood. Hence, there are reasons to believe that the uneven quality of TC live-action videos is in part due to the fact that TC video makers lack reliable and informed advice on what makes instructional live-action videos most useful for their intended users. “Standards” are not well established and “idiosyncratic application of design features” are common (Mogull, 2014, p. 340).

As a consequence, technical communicators therefore tend to adapt, knowingly or unknowingly, communication strategies that entirely depend on viewers’ immersion, the basic rationale being that if a procedure or process is clearly shown and/or demonstrated as it unfolds “live,” its users will understand it, if they watch/listen carefully. As this article will show, learning by immersion is indeed key but does not in itself warrant live-action instructional efficacy. The crux is that if designers of video instructions confuse procedures that are teachable with procedures that are not, what cognitive load theorist John Sweller (Sweller, Ayres, & Kalyuga, 2011) discusses in terms of the need to differentiate between biological primary information and biological secondary information in mediated instructional efforts, misguided instructional designs will be the result. Such designerly confusion, we believe, will aggravate the well-known usability-related problems associated with visual instructions that result in their being ignored by users (e.g., Daniel & Tversky, 2012; Ganier, 2004; Lundin, 2015; Novick & Ward, 2006; Söderlund & Lundin, 2017).

Related research on video efficacies

The research presented in this article has developed from a larger and evolving tradition of research on instructional videos. There is substantial evidence to indicate that the video format has the capacity to enhance content comprehension by reducing cognitive load (Boucheix & Forestier, 2017; Lowe & Schnoitz, 2008; Swenbring & Eriksson, 2017), motivating learning (Rieber, 1991), effectively communicating real-time reorientations (Tversky, Morrison, & Betrancourt, 2002), and, in line with this, supporting procedural knowledge and learning skills with respect to dynamic processes (Berney & Bétrancourt, 2016; Boucheix & Forestier, 2017; Boucheix & Schneider, 2009; Castro-Alonso, Ayres, & Paas, 2015; Höfler & Leutner, 2007; Lowe & Schnoitz, 2008; Tversky et al., 2002). There is also some evidence suggesting that live-action videos facilitate quick, “good enough,” procedural knowledge (Ganier & de Vries, 2016), in particular when live-action footage displays human hand movements and when the instruction communicates motor skills (Marcus et al., 2013). In the case of assembly instructional live-action videos, this is probably due to users’ preference for simultaneous watching/listening while assembling (Chirumalla et al., 2015).

Technical communication scholars, and Morain and Swarts (2012), embrace YouTube how-to videos as part of a “media ecology of tutorials” (p. 6). In their assessment of software video tutorials on YouTube, based on online user ratings, Morain and Swarts stress the priority of affective,
physical, and cognitive design aspects. Implicitly, they acknowledge the relevance of the long-since established cinematic conventions discussed in film production literature that basically serve to make complex content seamless, which, in turn, facilitate effortless watching, a topic further elaborated upon by Boucheix and Forestier (2017) and Swenberg and Eriksson (2017). In line with this, Morain and Swarts assert that one key affordance of tutorial videos is, what they call, viewability and link this to possible production constraints and designerly intentions. In accordance with Morain and Swarts, this affordance depends on, among other things, the video content’s resolution and poor or good “production” (p. 10). These are design aspects that belong to, what they label, videos’ physical design. Inferred from this, perceptual affordances may be defined as conditional, base-level aspects of an instructional medial object’s cognitive design that permit the instructions that are conveyed to become understandable.

Technical writing scholar Kara Poe Alexander (2013) recognizes usability-related advantages and limitations when employing “online video” for technical instructional purposes. In her printed text/video assessment comparison, she extends the many findings from multimedia educational psychology discourse into the TC domain. She concludes that the print version and the video version scored equally high in terms of their usability scores but also identifies the problem of users’ locating relevant information in the video version. This raises some key caveats with regard to the assumed efficacies of instructional transient media. These caveats might be discussed in terms of what Tversky et al. (2002) term the “congruence principle” and the “apprehension principle.”

The congruence principle states that effective graphic representations should externalize the desired content and structure of the internal representations that take shape in the minds of the representations’ users. Various types of abstractions are considered to trigger mental images most powerfully (Tversky et al., 2002). The implication of this principle is that live-action instructional format might be regarded as a medium that only fosters passive learning and/or superficial learning, because users do not need to mentally animate them (Clark & Mayer, 2016, p. 83). This sets the live-action format apart from, for example, abstract computer-generated animations used for training purposes (cf. Akinlofa, Holt, & Elyan, 2014). Live-action videos are, as opposed to animations, already “informationally complete” (Watson et al., 2010). This explains why James J. Gibson, the “father” of direct perception theory, when referring to celluloid film, noted that “Moviemakers are closer to life than picture makers” (1979, p. 182). When discussing the live-action format’s aspects of congruence, this also has bearing on the apprehension principle, that is, that the two principles are inter-related.

The apprehension principle states that its users should readily and accurately perceive the content and structure of the external representation. This explains the reason scholars of instructional media often discuss speed factors and the desired, but often missing, interactivity features. Motion might blur important details and does not easily invite reinspection (Leahy & Sweller, 2011; Paas, Van Gerven, & Wouters, 2007; Sweller et al., 2011; Wong, Leahy, Marcus, & Sweller, 2012). Motion also renders the creation of certain abstractions, such as making parts of objects semitransparent (as in structural diagrams), extremely cumbersome.

However, the photographic nature of live-action videos also presents some possible advantages for technical communicators. Because the live-action video formats normally do not require mental animation efforts on the part of viewers, they impose, in theory at least, low cognitive load in certain instructional circumstances. The implications of this are that some aspects of the overall pedagogical and instructional goal of the live-action instructional video that can be ascribed to a human, or groups of humans that display gestures, informal speech, descriptive movement, and so on, can be left without overtly attending to them. If so, more effort can be put into teaching about what is normally the primary goal of any instruction: the acquisition of biologically secondary knowledge. This is also to point out that previously acquired skills, biologically primary in nature, may leverage instructional attempts to promote acquiring biologically secondary skills (Paas & Sweller, 2012).

How well a particular instructional live-action video articulates the apprehension principle and congruence principle can be assessed indirectly by, for example, analyzing YouTube viewer-generated
data (Morain & Swarts, 2012), as well as usability tests, comprehension tests, and questionnaires (Alexander, 2013). However, a priori designerly assumptions about how to achieve such successful articulation is a whole other matter, because content and the expression of instructional live-action video communication generate their own layers of signification, what Kress and Van Leeuwen (2001) refer to as a multimodal “stratification” of meaning. The designerly implications of this stratification become especially critical to understand in the case of instructional live-action videomaking, whereby meaning potentials primarily materialize “live,” on-set, moments after the videographer has pressed the camera’s record button. In such instances, the strata of production and design become one (Kress & Van Leeuwen, 2001, p. 55). No live recordings are “picture perfect,” and this semantically blurred situation requires editing to make sure that the recorded content is consistent with the original instructional blueprint for design, and the stated skill acquisition end goal. It takes great skill to do this without compromising continuity, what cinema and eye-tracking scholar Swenberg, discusses in terms of achieving “perceptual precision” when editing live-action content (Swenberg & Eriksson, 2017), what Boucheix and Forestier (2017) label “the attentional continuity effect” (p. 360), and what Morain and Swarts (2012) address as a problem of structuring videos’ cognitive design.

### Multimedia guidelines as design approach

It is no secret that previous assessment studies that base their analyses on attentional engagement and various kinds of learning metrics are unsystematic and have generated inconclusive, and what can appear to be contradictory results (Boucheix & Forestier, 2017). A case in point, video content in this vein of research is regarded as the root cause for the increase of extraneous cognitive load (Clark & Mayer, 2016; Leahy & Sweller, 2011; Spinillo, 2011; Wong et al., 2012), and the decrease (Ayres, Marcus, Chan, & Qian, 2009; Boucheix & Forestier, 2017; Lowe & Schnotz, 2008; Wong et al., 2009). In part, such seemingly divergent results are due to the fact that scholars of video instructions study a wide range of different video formats that do not cognitively work in the same way (Höffler & Leutner, 2007). What adds to the confusion, is that researchers frequently adapt their own vague media format definitions that often are incompatible with the long-since established genres of film production/cinema theory. Most frequently used are the terms animations and video, but the terms dynamic, video-based, lecture video, transient, film clips, and multimedia are also used to describe various time-based medial formats that can be animations and live-action videos, and various hybrid forms of the two that may, or may not, include graphics, still pictures and, of course, audio (cf. Höffler & Leutner, 2007).

Yet inconclusive results and unintelligible definitions notwithstanding, generally, the cognitive turn in research on instructional “multimedia” has proven valuable to communicators and designers of audiovisual instructions, and there is substantial evidence of the validity of research-based design guidelines aimed at supporting attentional and psychological engagement. Over the last few decades, such guidelines have become prolific. Colin Ware’s (2013) guidelines amount to 170 examples (pp. 445–457), Sweller’s (Clark, Nguyen, & Sweller, 2006) guidelines can be summarized into nine primary examples, Mayer’s (Clark & Mayer, 2016) lists seven primary examples, whereas the design guidelines for tutorial videos of van der Meij and van der Meij (2013) amount to eight. Because human users of instructions all share the same basic human brain/mind architecture, design guidelines based on empirical findings extrapolated from scientific experiments concerning the relation between visual (or audiovisual) stimuli and the human brain/mind architecture also should have significant bearing on designerly matters related to the instructional live-action genre.

However, questions arise as to what design guidelines are relevant in what medial and communication contexts, what are the relevant boundary conditions for live-action videomaking? As of yet, these questions remain unanswered to some extent. This situation generates constraints with regard to the sought-for transferability of results, for instance, from the domain of visual instructions to the domain of technical documentation (Große, Jungmann, & Drechsler, 2015), and, more generally, for further research into how cognitive engagement relates to issues of usability, research that is
necessary for providing evidence for the quality of instructional tools (Henrie, Halverson, & Graham, 2015).

**HCD as design approach**

Users’ support tool preferences and their “cognitive styles” (Höffler, Koć-Januchta, & Leutner, 2017) vary. Consequently, it is a tall order to achieve optimal usability for all kinds of learners when designing TC videos. Therefore, it is wise to remember that the notion of “good enough” is at the heart of all worthwhile design endeavors. Although the term *good enough* perhaps connotes a self-defeating designerly stance, prominent design scholars Simon (1975), Cross (1982), and Lawson (2006) argue, in different ways, that settling for the good enough should be considered the best designerly outcome. If good-enough solutions are to be obtained, what Simon (1975) discusses in terms of satisficing instead of optimizing, a broad range of alternative design outcomes and solutions must be identified. According to Cross (1982), they have to be “actively constructed by the designer’s own efforts” (p. 224). This translates into a particular mind-set—designerly ways of knowing—which entails various kinds of reasoning activities that reframe the problem space primarily by synthesis, and not by mere analysis (Cross, 1982).

Following this line of reasoning, the HCD champion Krippendorff (2006) argues that this reframing can only be realized if transformation and meanings are “recursively woven into interaction with others” (. 70). This, according to Krippendorff (2006), is a “second-order-understanding” that is dialogical in nature and recognizes that, because meanings ultimately depend on the perceptions of consumers and end users, designers must interact with them and engage in designerly conversations. This line of thought provides a window into the proposed virtues of HCD and the contemporary form of HCD (it has indeed a rather long history not accounted for here) that represents a deemphasis on the view that humans primarily are cognitive information–processing organisms.

Although live-action instructional videos per se are not a focal point of HCD researchers, visualization technologies and information/knowledge support systems are. In this context, communication objects’ cognitive design are approached in terms of factors that condition usability. In the field of TC, according to Zachry and Spyridakis (2016), HCD-inspired research means that technical communicators are encouraged to assume a stance that acknowledges that visualization technologies and information/knowledge support systems have sociocultural implications that reverberate further than the rather uncontroversial and confined industrial context, and that TC scholars should not shy away from provocative research-based arguments.

In spite of the fact that HCD when approached as a form of theoretical framework, appears to be philosophical rather than descriptively instrumental, there are some good examples of how HCD-inspired ideas can be put into practice. The HCD-inspired study by Kluge and Termer (2017) highlights the design process of an application for mobile devices. The researchers claim there is a HCD methodology and that this methodology was employed throughout the design process that proceeded in four distinct steps (pp. 172, 179), described as follows:

1. In-depth interviews with end-users (containing three primary questions: who will use the application, what they will use it for, and under what conditions will they use it?).
2. Development of fault scenarios
3. Layout and concept design development
4. Evaluation and testing in the field.

Kluge and Termer (2017) suggest that the so-called HCD methodology aided in maximizing technology acceptance and conclude that the importance of the HCD approach during the development process was critical in the attainment of a deeper understanding of the domain within which problem solving needed to be supported. They further stress that only if relatively low levels of user
acceptance, technology acceptance, and job satisfaction are achieved, there is a definite added value of using new technology (p. 179).

Along similar lines, maritime technology scholars Costa, Holder, and MacKinnon (2017) report from a visual display design project that employed the ISO 9241-210:2010 for HCD for interactive systems. They describe HCD knowledge transfer and integration in terms of various actor activities, that also include nonhumans (e.g., HCD literature, project plans, screens, etc.), by identifying connections between actants and how they mobilize knowledge transfer. The researchers consider the transformation of what they call HCD theory into industrial practice as a series of translations and negotiations. Overall, the theoretical discussion of Costa et al. exemplifies the designerly urge to, on the one hand, regard HCD as a philosophy, inviting the design team to seek out new solutions to an open ended problem, and, on the other hand, as a kind of design process model with clearly defined stages and objectives. They consider the latter approach as a problematic one.

The previous TC research above are examples of how HCD-inspired research has the capacity to illuminate sociocultural factors that influence the usability of a particular design. In the case of visual instructions this is key, because biologically secondary information that is culturally acquired knowledge is the primary concern of designers of instructions (Sweller et al., 2011). In other words, HCD-inspired instructional assessment studies fill an important knowledge gap with respect to how video functions as an information and knowledge support tool in real-life communication scenarios.

**Perceptual affordances of the live-action format**

As highlighted in the previous HCD theory section, Krippendorff (2006) considers scholarly design problem-reframing activities that are dialogical in nature a second-order understanding. With regards to designerly matters that pertain to TC videos of the live-action format, what, then, might represent a first-order understanding? In this article this first-order understanding is generated by employing affordance theory.

The plethora of psychology, design, and communication research efforts that use affordance theory is indicative not only of the multifaceted nature of affordances but also of its scientific usefulness when assessing communications (Evans, Pearce, Vitak, & Treem, 2017; Rice et al., 2017). However, in this article it is especially appropriate to use the concept of affordances the way Gibson (1979) originally envisioned it because it was developed out of Gibson’s work with instructional films. During World War II, James J. Gibson was in charge of assessing instructional films for pilots as a lieutenant colonel in the U.S. Air Force. Therefore he had rather solid empirical reasons for his understanding of how the film medium faithfully captures essential features of reality that may enable humans to act upon them in purposeful ways. Allegedly, this experience informed the (later) formulation of the tenants of the ecological perception theory that is based on the notion that human vision is a flowing optic array, oriented toward the horizon (Bordwell, 2005; Gibson, 1947; Reed, 1988).

For the most part, in the field of technical and professional communication (TPC), the basic understanding of affordance appears to come from Kress (2010) and the social semiotic approach to contemporary communication or/and the way it was popularized by Norman (1999) with regards to notions of computer interface usability. Contrary to Kress and Norman, Gibson (1979) argues that the gateway to the environment is in the form of actual information provided by arrays of energy that are patterned and quantifiable. This information—perceived directly—enables humans (and other animals with camera eyes and ears) to actively listen and look for the things that the environment can afford us. In other words, actants directly perceive the opportunity for action within the arrangement of the environment. This idea is related to Gibson’s (1979) view that meaning is not a cognitive act and is already “present and available to actors in the environment” (Beynon-Davies & Lederman, 2017, p. 145). Hence, affordance theory, in concurrence with how Gibson originally envisioned it, is all about the picking up of cues via the sensory systems over extremely short time spans (as brief as 300 ms.).
For scholars of visual/audiovisual instructions, Gibson’s affordance concept is therefore well suited to explore human actions in relation to material artifacts used for informative purposes as a “first-order” connection between articulation and communication (Beynon-Davies & Lederman, 2017).

Following this argument, it is feasible to develop a scheme of classification of live-action video affordances, that is, a kind of taxonomy of perceptual affordances (cf. Teston, 2012). Informed by film/video production value/image quality guidelines (Eriksson & Eriksson, 2015), what Morain and Swarts (2012) discuss in terms of primary traits and physical design aspects of instructional videos and the set of criteria offered by computer-mediated communication scholars Evans et al. (2017), we suggest that live-action video imagery primarily embodies four key perceptual affordances: comparability, verifiability, recordability, and visibility.

**Visibility**

The most important affordance of the live-action format might very well be considered visibility, or what we might label as, for lack of a better term, audiovisualability. Although not unique to photographic genres, it is beyond doubt that without visibility, the other key affordances would be severely weakened, which is often the case in amateur video productions, due to a number of reasons, including poor resolution, shaky camera handling, poor dynamic contrast, poor lighting conditions, jump cut edits, dirt on lens, etc. (cf. Morain & Swarts, 2012). In other words, without high visibility, no other affordances may be fully exploited (nor analyzed). This also implies that the size and quality of screens when watching to/interacting with a live-action video probably matter, as larger screens afford greater visibility, and greater visibility makes the discerning of great amount of detail possible (Leonardi, 2013). In any case, visibility depends on to what degree something is recordable.

**Recordability**

What really sets the live-action format apart from static photographic genres is the affordance of recordability. This affordance highlights the videographer’s status as designer, for instance, the ability to evoke sensations that are linked to human ambulatory vision. Recordability also highlights the live-action format’s propensity to reveal irrelevant, visual, and audiovisual elements. The world as we know is not always picture/audio perfect. Hence, recordability is a boundary constraint. Not all things are recordable, though we sometimes wish they were. For example, objects’ interiors are not. If professionally exploited by a videographer, recordability also makes possible the activity of editing in postproduction. This gets to the videographer’s status as a director/designer and his or her hopefully well motivated selections of what content to record (and what not to record) to be able to in the editing phase make the communication of instructions more brief, more to the point, less meandering, and, hopefully, less boring. This affordance is also, it appears, linked to immersion learning strategies for similar reasons as verifiability but emphasizes the role of recording technologies that imprint the recordings in different ways, for good and for worse (Eriksson & Eriksson, 2015).

Verifiability affords an accessible means of environment identification. It is likely that without verifiability, live-action video instructions would not function very well in immersion learning efforts because learning by immersion appears to be intrinsically tied to extremely effortless recognition (Paas & Sweller, 2012). The indexical quality of photographic images cannot be underestimated, though, according to cinema and documentary scholar, Bill Nichols (2010), it is subject to qualification since photographic images cannot show everything, as they can be altered (and often are) and do “not guarantee the validity of larger claims made about what the image represents or means” (p. 42). If images embody high degrees of verifiability, they are easily used for the activity of comparison, albeit “fake” live-action imagery also function in this regard.
Comparability

Generally speaking, the affordance of comparability can be said to be of high relevance when procedures in how-to videos are being mimicked in actual, hands-on, learning situations. Comparability affords the activation of humans’ mirror neurons. This is, so to speak, comparability at work in the mind and highlights the fact that action execution and observation are intimately related processes. One practical implication of this is that live-action video instructions should facilitate users’ (almost) simultaneous action and watching for users to fully exploit the mirror neuron activation advantage (Chirumalla et al., 2015). This has profound implications for learning because compare and contrast is our first and foremost natural form of thought (Silver, Strong, & Perini, 2007). This affordance also facilitates simultaneous intramedia comparisons to various degrees, for example, audio track with video track in the same instance. Objects should sound the way they look, and vice versa. Recordability is thus enhanced if the videographer also record, synchronous, high quality audio in addition to the visuals (cf. Morain & Swarts, 2012).

Here it is important to note that the aforementioned definitions of key live-action video affordances do not exclude other key (perceptual and other) affordances. Indeed, no measurements and data categories, no matter how refined and intelligible they are, may encompass and make all potential affordances that are relatable to live-action instructional videos and their ontologies identifiable. It is very unlikely that there ever will be one final exhaustive set of live-action video affordances (cf. Rice et al., 2017). Indeed, affordances tend to be sublime and may be considered so “instinctual” and “routinized” (Burlamaqui & Dong, 2016) that they, depending on who perceives them, easily become “invisible” (Ortmann & Kuhn, 2010). As noted by Stendal, Thapa, and Lanamaki (2016), it is a great challenge to identify all affordances, categorize and label them. Moreover, on epistemological grounds, some scholars reject the notion that affordances can be evaluated at all or and made distinct from other environment–user relations (Bygstad, Munkvold, & Volkoff, 2016; Rice et al., 2017).

Contribution and research questions

In short, research on design of live-action instructional videos is unsystematic and investigations into production matters related to videography has mainly been a non-academic concern (e.g., How To Shoot Video That Doesn’t Suck by Steve Stockman, 2011). The overall aim of this study, therefore, is to develop a clearer understanding of the affordances of live-action instructional videos, how these affordances relate to usability and designerly ways to support live-action usability in actual video production. In this article, we do this by exploring a more intersecting and explorative assessment perspective that combines aspects of cognitive load and direct perception theorizing with one that serves to acknowledge stakeholders’ critical roles (Krippendorff & Butter, 2007).

Here, it is important to note that the term usability is used in a rather straightforward manner: if a live-action instructional video can be said to fulfill its purported function, then it exhibits usability (cf. Erlhoff & Marshall, 2008; Rouse, 2007).

The objective of this study is to identify links between the basic instructional design strategies of the video that is the focal point of this article, and how it is perceived by several groups of end users over the course of a longitudinal design process. Following this, the analysis of this article revolves around a case study of a live-action instructional video design project and is driven by two questions that pertain to perceptual affordances:

Research Question 1: In what ways does the live-action video affordances make possible its desired, instructional, expressive capacity?

Research Question 2: In what ways does the live-action video medium affordances constrain or resist its desired, instructional expressive capacity?
Design process and methods

In this method section, we will present the mapping of the cognitive design events of the instructional live-action video that is the object of study of this article and its associated design research framework. This design research framework is inspired by RTD. This translates into a professional stance with regards to design endeavors. Here this means that the authors of this article also are designers (of the video support tool in question). This is to acknowledge that complex designs activities require explicit knowledge and tacit knowledge and that designers with ample and relevant design related experiences may better make sense of both of these types of knowledge (Löwgren, 2016). More specifically, the primary reason for maintaining creative control in this way is because we think intentionality is an aspect that should not be taken lightly when analyzing and evaluating a medial object. If the rationale behind an erroneous design choice is largely unknown, it is hard to judge whether a proposed alternative approach would really generate better results.

This case study’s research framework consists of the production stages with assessment activities (interviews, usability reports, and YouTube user generated viewing data). The different research methods were chosen in a systematic fashion and used progressively to clarify the objective and problem statement, and to make a better TC video. In other words, this design research project bears the hallmarks of a design research project in which the designed object is expected to become more and more refined as its users provide feedback. We use design research methodology (DRM) (Blessing & Chakrabarti, 2009) as our design process framework. DRM is well suited for the development of support tools (p. 143). DRM allows for iterative assessment activities and provides design researchers with a stepwise, goal-directed, design research framework that consists of four stages: a literature review, task clarification, and formulation of research goal (research clarification); empirical data analysis and initial description of existing situation (descriptive study 1); identification of desired situation and improvement of problem definition (prescriptive study); and, lastly, an investigation of the communication impact of the video (descriptive study 2). Albeit limited in scope, the design process journey of the video in question represents all of the DRM research stages, See Figure 1.

Materials: the video

The assembly/construction methods the video conveys are displayed by a team of antipredator fence builders. The video thus demonstrates, and partly documents, how such an antipredator fence ideally should be constructed. The video’s target audience, that is, its usership, are Swedish farmers and cattle owners who wish to protect their livestock (primarily sheep and cows) from wolves and bears

---

**Figure 1.** A stylized image showing the design research process of the video.
and, by doing so, would be eligible for an European Union (EU)-financed government subsidy that will cover the cost of fencing material, if accurately built in accordance with certain quality standards. Hence, the overall aim of the video is to enable users to construct a robust, high-quality fence that after inspection by government officials would be eligible for the subsidy. The final version/prototype (Prototype 2), available on YouTube, is 27 minutes long.1 It is titled “StängselManual.” Its verbal instructions, as well as added text, are all in Swedish. See Figure 2 for a still from video’s Chapter 3.

The video is the audiovisual result of a collaboration between the authors of this article and a Swedish government agency (Länsstyrelsen, Gävleborg), a Swedish wildlife conservation organization (Rovdjursföreningen), a Swedish educational organization (Studieföreningen), as well as several companies that manufacture and/or sell fencing equipment (Swedol, Granngården, XL-Bygg, Gunnebo, Fyrås Trä, Bole, and Delka produkter AB). The video was shot and edited by one of the authors of this article (Per Erik Eriksson, here referred to as the “videographer”). The videographer is a former TV producer, professional videographer, and video editor.

Research clarification: identifying the overall instructional design strategy

The overall design of the video was informed by the basic notion that the video ought to communicate to livestock owners the best and most expedient way to construct a high-quality antipredator fence. This would be to exploit the potential to foster stakeholders’ acceptance of a useful video as a solution to a problem. Predator attacks on livestock is of great concern for many sheep and cattle owners in central and southern parts of Sweden and is a widely debated problem. It is common that farmers resort to controversial (and illegal) solutions to this problem. In this situation, better livestock fences would serve all parties well. At this point in time there were eight other (Swedish) fence construction videos available on YouTube. Each of them had been viewed a few hundred times or less, they drew overall poor ratings and bore the hallmarks of what Morain and Swarts (2012) label “poor videos.”

Figure 2. A screen-shot from the antipredator fence instructional video, showing the fence-building instructor in action as he shows how a “parallelogram” is constructed. To the right, macroscaffolding (“Chapters”) in the form of text and graphics are visible.
To be eligible for the government subsidiary, the antipredator fence in question should remain standing for at least 20 years and must remain 100% functional for at least 5 years. Therefore, quality was identified as the most important aspect of the preliminary set of success criteria against which to evaluate the outcome of the research (Blessing & Chakrabarti, 2009, p. 43). It was then speculated that this preliminary set of success criteria would, if adeptly communicated, result in another, secondary, success criteria, which is many views on YouTube. At least significantly more views than other similar videos that appeared to be underutilized.

The fence itself and the construction thereof should therefore represent an ideal result. In instructional terms, this means that the instruction must facilitate users’ accurate understanding of how to construct such a fence, not only what the fence should look like, what it consists of, the cost, etc. In other words, the instruction ought to work at two levels, the “how level” and the “why level” (Eriksson, 2017). This is the equivalent to what TC scholar J. Swarts calls “demonstrative content” (2012). A good video instruction that people enjoy watching should focus equally on doing and explaining (Swarts, 2012, pp. 195, 198). The why level concerns content that aims to explain/clarify why a certain action is executed and entails relational and interconnectivity aspects. The how level, on the other hand, is concerned with how to do things and involves distinct actions that are shown to enable the user/viewer to imitate/act upon certain key actions. It was expected that most of the longer sequences of the video would internalize both of these communicative aims.

For the construction of the fence to be understandable, it was decided in collaboration with the Länsstyrelsen government agency that the video should be structured in accordance with six construction key themes: workflow, posts, parallelogram, isolation, power, and connection/support materials. Together, these themes communicate all necessary facts and other relevant information in the form of a narrative, an ideal construction procedure, from start to finish. This forms the basis for the video’s cognitive design (cf. Morain & Swarts, 2012), see Figure 3.

**Descriptive study: producing/designing video prototype 1**

Before the actual recording of the video, the expert fence builder (*Delka produkter Inc.*) and the videographer discussed the overall idea about an instructor’s presence (cf. Wang & Antonenko, 2017) and the pedagogical goals of the video: (1) inform objective, (2) stimulate recall of prior learning and experiences, (3) present learning materials, (4) activate responses (from the other

![Figure 3](image-url). A stylized image showing that the video’s content is structured in accordance with key themes and their associated communication ideals.
characters in the video), (5) provide feedback (to fence constructor team members visible in the video), and (6) assess the performance.

It was decided that the videographer would accomplish this by talking directly into the camera at times, as well as by prompting the relevant talking points. In this way, the expert fence builder would be able to focus on the actual construction of the fence and not be hampered by trying to remember the pedagogical goals. The first version of the video was shot and edited to our best abilities as technical communicator video makers, it was consistent with several key multimedia design guidelines (such as guidelines that effectuate the modality effect) and consistent with cinematic production standards in terms of lighting, audio, camera angles, camera movement, framing, exposure, focus and so on. The first video prototype running time was 12 minutes. It was recorded on an HVX-200 Panasonic HD camera (720p).

**Descriptive study: focus group interview**

Video Prototype 1 was evaluated in screenings that consisted of five presumed video users, all of whom are keepers of livestock in a rural part of Sweden (in an area with wolves and bears), and with prior experience of electrical fence construction. All the livestock keepers run independent farming companies. Three individuals are male, and two are female, with an average overall age of 48 years. The interview was in the form of a probing, open-ended, group interview lasting for 1 hour 45 mins. During the interview, sequences of special interest were rescreened and further elaborated upon. Three primary questions that revolved around the video’s cognitive design were addressed: (1) What sequences are perceived as presenting facts and actions in an accurate way (if any), which ones are not (if any)? (2) Is the structure of the video understandable; do the events/actions portrayed appear complete, can they be applied? (3) Is the content pertinent; does it have an instructional purpose, if not, why? The answers of the presumed users were summarized and extrapolated into a checklist of the video’s six sequences to see which actions they had (fully) grasped, to identify possible ways to improve the video.

Potential improvement schemes were then, in turn, identified by relating them to the key classification themes: comparability, verifiability, recordability, and visibility. This was accomplished via the means of coding the focus-group transcript, that is, making manifest particular content that best demonstrate “latent” but relevant verbal data that include perceptual and videography related matters. For example, when the participants talk about something in the video being too small, this is classified as a problem of visibility. Comments that concern things being out of sight, on the other hand, is identified as a concern that fits within several classification themes (for example, visibility, recordability and verifiability). In other words, some classification themes are allowed to overlap, some are separate.

**Prescriptive study: video prototype 1.1 and the in-field try-out**

In the prescriptive study stage the video was further evaluated in an in-field try-out. At this stage, in accordance with the recommendation of government agency Länsstyrelsen, two previously not included sequences were added, one “talking head” sequence at the beginning of the video, with a government official explaining how to apply for the government fence construction subsidy, and one at the end of the video about how and why one might construct an alternative antipredator fence with a welded sheep wire fence. This resulted in the video being 19 mins. long. The video’s usability was then evaluated by two fence construction teams in an in-field try-out. The two teams were considered high ability, in terms of electrical fence construction. The team managers of both teams, those who primarily used the video support in question, are two husband and wife couples who run two separate sheep-farming companies. The team managers of Team 1 had an average age of 35 years, whereas the team managers of Team 2 had an average age of 62 years. The teams’ feedback
consisted of written responses, which were further elaborated upon over the phone shortly after the responses were received.

The teams were urged to write the responses immediately after using the video in the field and were instructed to elaborate upon three inter-related questions: (1) To what extent the information presented in the video helped them to construct the fence, (2) To what extent the content of the video was easily understood and applied, and (3) To what extent the video supported their knowledge and information needs during construction. The transcripts of the post self-reports conversations as well as the self-reports were coded (see previous description of how this was operationalized).

**Descriptive study 2: prototype 2 and Youtube analytics data**

The cognitive design and usability assessments of descriptive study and prescriptive study resulted in reedits, rerecordings, and additions of graphics (see “Design Measure” in Results section). The additions of graphics, when applicable, were made in accordance with relevant design guidelines. One of the most obvious changes to the video was the addition of a graphic representation of a chapter index that was made visible on the right-hand side of the video frame. This index contains chapter “boxes” that sometimes become red, one box at a time, indicating for the viewer how far the video’s narrative has proceeded. According to Cojean and Jamet (2017), this is an example of macroscaffolding that promotes information-seeking behavior. This design addition is also an example of the application of the segmenting and pretraining principles (Clark & Mayer, 2016, pp. 201–218). The Prototype 2 extended to 27 minutes in length. After the Prototype 2 of the video had been viewable on YouTube for an extended period of time, we used YouTube analytics to further assess the video’s usability.

**Results**

A selection of basic YouTube user-data show, as of March 2018, that Video Prototype 2 has generated 8,270 views. The average viewing time is 5:43 mins. Viewers’ gender are male 81%, female 19%. Viewers’ ages are 13 to 17 years 2%, 18 to 24 years 6.5%, 25 to 34 years 29%, 35 to 44 years 23%, 45 to 54 years 18%, 55 to 64 years 14%, and older than 65 years 8.6%. Screening technologies are computer 52%, smartphone 29%, other handheld devices (such as the iPad) 17%, and others/unknown 2%. Average viewing time/screening technology is computer 6:00 mins, smartphone 4:51 mins., and other handheld devices (such as the iPad) 7:00 mins.

In what follows, we present the results of the users’ assessments of the video: the focus group interview (see Table 1), the in-field try-out (see Table 2), and YouTube relative attention user data (see Figure 4).

**Discussion**

The results of the case-study presented in this article show that various kinds of audiovisual deficiencies if identified in a concerted and structured fashion drive design improvement schemes. Such transformations would have been difficult to accomplish if the video’s users were not consulted to identify possible validity threats to the instructional design, in this case, factors that diminish the “strength” of the perceptual affordances of the live-action video format. Design process management frameworks, such as, for example, the DRM in combination with HCD-inspired approaches and RTD approaches offer constructive ways to do this. The downside of this overall design approach if actually applied in a conventional technical video production setting is that it renders, improvised, “run and gun,” quick turn-around, video production impossible. However, the video production case study featured in this article is a video support research project that primarily serves to identify
Table 1. Table shows video sequences’ content and the associated focus group’s feedback, and, in turn, how this relates to live-action affordances and possible design measures.

<table>
<thead>
<tr>
<th>Description of Video Sequence/Actions</th>
<th>Summaries of Focus Group’s Feedback</th>
<th>Live-action Affordance</th>
<th>Design Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workflow: Expert talks into the camera, describes how and in what order key actions should be executed.</td>
<td>No comment.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fence Posts: Sequence shows expert and construction team fastening the posts. Expert and videographer talk about the reasons why a sturdy fence consists of straight lines. Expert gives team feedback on how to make them stand up straight. Sequence also shows ideal digging and refilling methods. Expert gives feedback on pole’s position in hole and the depth of the hole.</td>
<td>Key information with respect to sizes and distances is missing. Understands extent of force required. Understands how much time is required. Understands position of tool and workers. Inferred differences of refilling methods unclear. Yardstick used in action provides key information, but numbers on yardstick are small.</td>
<td>Comparability - No way of seeing/understanding distances and sizes, comparisons cannot be made. Visibility/Verifiability/Recordability - Clearly sees bodily effort and facial gestures that communicate exertion. Sees/hears how force gradually drives posts downwards, understands extent of workload. Visibility - Key materials and tools not clearly visible. Verifiability/Visibility/Recordability/Comparability - Yardstick provides approximate information about sizes and distances. Verifiability/Visibility/Recordability - Key explanatory soundbites are missing in recorded video. Camera movement - a pan - audiovisually. Increased tension is verified visually. Gripple fits into overall construction.</td>
<td>Reedit sequence. If possible, insert footage that shows distances and sizes more clearly. Explanatory cues and graphics may also be added.</td>
</tr>
<tr>
<td>Parallelogram: Expert describes what a “parallelogram” consists of. Expert also describes the placement of horizontal pole that connects vertical poles. Expert and assistant show how the vertical pole is connected using a chain saw. Expert secures pole with two nails. Expert fastens the parallel wire in the shape of a number eight, explains why this should be, and why it needs to be attached to the vertical poles. Expert inserts wire into Gripple and tightens wire by hand. Expert tightens the gripper with the Gripple Torq Wire Tensioning Tool and explains why it is used.</td>
<td>Expert fails to motivate in a coherent way why a parallelogram is needed. Impressed by the fact that the post is fastened in a nail-free fashion (nails only secure pole’s position), this is motivational. Do not see the “eight shape”, do not fully understand the reference to the “eight shape”. Uttering with explanatory sentence is unfinished. Understands how Gripple functions and understands how tension is regulated by the tensioning Tool. Understands how the Gripple fits into overall construction.</td>
<td>Comparability - Discrepancy between audio and video track - video shows shape, audio is inconsistent with this. Verifiability/Visibility - Lack of perspective/illusion of 3D in combination with relative smallness of wire render meaning of shape-metaphor unclear. Gripple’s interaction with wire is visible, wire’s increased tension is verified audiovisually. Comparability - Audio information can be compared with visual clues in video footage. Comparability - No way of seeing/understanding distances and sizes, comparisons cannot be made. Visibility/Verifiability/Recordability - Key explanatory soundbites are missing in recorded video. Camera movement - a pan - audiovisually. Increased tension is verified visually. Gripple fits into overall construction.</td>
<td>Reedit sequence. If possible, insert footage that includes introductory statements about the specific function of the parallelogram in the overall system. If possible, find footage that shows the “eight shape,” combine this with better explanatory audio. Locate and include close-up that displays wire. Explanatory cues and graphics may also be added.</td>
</tr>
<tr>
<td>Isolation: Expert explains the function of insulators and why certain insulators are better than others. Together with a fence construction member, the expert shows how and where (on the post) the insulator is placed.</td>
<td>Alternative, potentially valid methods are mentioned, but not explained, nor shown. Understands required workload and placement of insulators, but is not convinced of the superiority of the material chosen.</td>
<td>Comparability - Markings on insulator placement tool are visually comparable. Visibility/Verifiability/Recordability - Sees/hears how relatively little force is required to drive in nails.</td>
<td>Record and add sequence with other isolation materials that are of the how/why kind.</td>
</tr>
<tr>
<td>Power: Expert explains how electricity gets transmitted in all directions and how the predatory animal receives an electrical shock.</td>
<td>Keep it short, no need to expand on this topic. Electricity and circuits are key, but not in this video.</td>
<td></td>
<td>Shorten sequence to make it less redundant.</td>
</tr>
</tbody>
</table>

(Continued)
audiovisual communication deficiencies, not to produce a most exemplary video with optimal usability in a short amount of time with limited recourses.

With the knowledge that viewing statistics on YouTube certainly are no perfect way to gauge communication efficacies, the facts that the video in question has been viewed more than 8,000 times (as of March 2018), and the overall positive general comments of the in-field try-out teams, indicate that the video that is the focal point of this article probably embodies usability aspects that other similar Swedish livestock fence construction how-to videos lack. In any case, other such videos have significantly fewer views. This general argument is reinforced by comments from users on YouTube (in the “Comments” field), for instance, from Daniel Andersson (translation from Swedish), “I have watched the film multiple times, because I am constructing such a fence myself. Eventually, lots of details become visible, details that are not discerned at the outset.” This comment from Andersson indicates that the video is not only useful for him—for one reason or another—but also appears to be informationally dense. Apparently, some of its information needs to be revisited repeatedly for the content to be fully grasped. This may come as no surprise because user feedback in early product development tends to result in designerly additions of content, as opposed to deletions. It is an often-occurring problem in the product development context that what users want is different from what they can cope with. This information richness conditions the live-action format’s immediacy, what Rouse (2007) considers the urgent aspect of usability.

The problem of information overload gets to the heart of live-action, how-to video communication. This video design research project illuminates particular formations of interface information overload that, we suspect, significantly contributes to the decrease of users’ attention levels or, at the very least, people’s willingness to spend time watching certain parts of the video. This might explain why the average viewing time is slightly less than 6 mins., out of 27 mins. (however, this is a little more than a minute longer than the average viewing time for videos on YouTube). We suggest that these formations primarily are the unfortunate side effect of the application of certain conventional video production design strategies, that is, edits that serve to compress time and eliminate redundancies. For example, if information that exists in between lines of explanatory audio is omitted in editing, such as the display of bodily movement and diegetic/synchronous audio, this may result in that comparability and verifiability are being compromised.

Such edits, in of themselves, increase the prospect of causing discontinuity that, in turn, increases cognitive load, unless executed with great precision (Swenberg & Eriksson, 2017). Enhanced videos (e.g., Lamb, 2015) that embody techniques that make transient information denser and more concise do not necessarily augment instructional videos’ efficacies. It might seem that another possible explanation for relatively low attention levels is the difficulty of the associated material to be learned. However, the relationship between task difficulty and attention is far from obvious (Sweller et al., 2011).

<table>
<thead>
<tr>
<th>Description of Video Sequence/Actions</th>
<th>Summaries of Focus Group’s Feedback</th>
<th>Live-action Affordance</th>
<th>Design Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connection/Support</td>
<td>Does not fully understand how</td>
<td>Comparability - Imperfect</td>
<td>Locate and include close-ups</td>
</tr>
<tr>
<td>Materials: Expert explains why post</td>
<td>to tie one of the knots.</td>
<td>knot’s untying can be</td>
<td>that display the wire-tying</td>
</tr>
<tr>
<td>support is required. Shows how one</td>
<td>Understands why they need to</td>
<td>compared with expert’s</td>
<td>procedure, insulators and</td>
</tr>
<tr>
<td>kind of post support is</td>
<td>be tied that way. Does not see</td>
<td>explanation.</td>
<td>conductors. Explanatory</td>
</tr>
<tr>
<td>constructed with a chainsaw.</td>
<td>isolation in spite of what it is</td>
<td>Visibility/Verifiability/</td>
<td>cues and graphics may be</td>
</tr>
<tr>
<td>Shows it with a nail. Expert shows</td>
<td>referenced to.</td>
<td>Recordability - Sees how</td>
<td>added.</td>
</tr>
<tr>
<td>how to tie two different wire knots.</td>
<td></td>
<td>imperfect knot gets untied.</td>
<td></td>
</tr>
<tr>
<td>Explains why mere twisting will not</td>
<td></td>
<td>Sees stiffness of knot.</td>
<td></td>
</tr>
<tr>
<td>suffice.</td>
<td></td>
<td>Isolation has poor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>visibility. Camera fails</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>to reveal complete</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>sequence of knot-tying.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>initial bending sequence</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>is not visible.</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. (Continued).
Table 2. Summaries of fence construction teams’ feedback from the in-field try-out, some of which relate to specific video sequences. In the table, this is in turn related to certain live-action affordances and possible design measures.

<table>
<thead>
<tr>
<th>Summaries of Feedback of Fence Construction Team 1 (General and sequence specific)</th>
<th>Summaries of Feedback of Fence Construction Team 2 (General and sequence specific)</th>
<th>Live-action Affordance</th>
<th>Design Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>General comments: Overall video was helpful. Mostly easy to copy procedures, easy to follow, video is well-structured, but difficult to revisit parts of sequences - do not know how far along video has played. Rewind function works poorly in the field. Materials and tools are clearly shown. Bad weather (in video) is discouraging. Video does not include methods that pertain to construction of fence at waterways and ditches. Choice of materials imply that video is probably geared toward sheep owners, as opposed to cattle owners. Understands necessity of careful execution for the fence to last. Watching on a bigger screen (bigger than a normal smartphone) works best. Expert talks too much at times.</td>
<td>General comments: Video is useful, it is well-structured, but probably works better for novices. Video’s instructional content is too circumscribed, (our) previous knowledge not fully acknowledged. Some techniques are completely new for us and require more thorough instructions. In general, there is too much talk, too little doing and showing. One chapter heading/rubric (graphics) includes a spelling error. Video does not include methods that pertain to making fences work properly in snowy/icy conditions – which it should. Ideal thickness of wire is not communicated.</td>
<td>Visibility - Some actions and materials are visible, while others are not. Verifiability - Degree of force, exertion, complexity and materials’ ontological dimensions can be verified audiovisually. Comparability/Verifiability - Interplay between visual and audio stimuli makes certain comparisons possible. Previously known methods, sizes and dimensions can be compared with new ones and, therefore, verified. Others cannot. Recordability - Some unwanted phenomenon/elements (such as bad weather, unwarranted silence or odd utterings) become prominent.</td>
<td>Re-edit sequences or re-shoot material that does not exist in the original recordings. Add visual cues and explanatory graphics. Add graphics that allow for greater user interactivity that permits revisits and status of procedural content.</td>
</tr>
<tr>
<td>Subsidy: No comment.</td>
<td>Subsidy: Government official in the video (in the “subsidy sequence”) is no longer a government official, but works elsewhere now.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parallelogram: Shape of wire in parallelogram is unclear, requires multiple re-screenings.</td>
<td>Parallelogram: Safety concerns when handling chain saw are not touched upon</td>
<td>Verifiability/Visibility - Lack of perspective/illusion of 3-D in combination with relative smallness of wire render meaning of shape-metaphor unclear. Verifiability/Comparability - Preconceptions about safe chainsaw handling are inconsistent with actions portrayed in the video. Comparability/Visibility/Verifiability - Placement of insulator provides critical information about other materials that are not visible. Visibility/Verifiability - Video does not show and/or address ideal distance to power lines.</td>
<td>Re-edit sequence. If possible, insert footage that shows the “eight shape” more clearly.</td>
</tr>
<tr>
<td>Isolation: No specific feedback</td>
<td>Isolation: Placement of insulators reveals ideal position of wire - but the sequence is not about the wires, which is confusing.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Power: Perhaps needs to be discussed/shown in a separate video. Topic too complex.</td>
<td>Power: What about safety distances to power lines?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
So it seems that some instructional approaches harmonize with live-action video how-to communication, and some do not. A case in point are the video sequences that are associated with the lowest attentional levels, the Isolation and Power sequences. In spite of one of these sequences

![Figure 4](image)

**Figure 4.** Figure shows the YouTube analytics’ viewers’ relative attention graph. The figure consists of a screen-shot from the YouTube analytics page (in Swedish) and, below, added stills from some of the video’s sequences.

---

**Table 2.** (Continued).

<table>
<thead>
<tr>
<th>Summaries of Feedback of Fence Construction Team 1 (General and sequence specific)</th>
<th>Summaries of Feedback of Fence Construction Team 2 (General and sequence specific)</th>
<th>Live-action Affordance</th>
<th>Design Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connection/Support Materials: Gallagher-knot difficult to discern (on a smartphone). No idea tying, as very strenuous (and suitable only for very strong hands), this is not communicated. Sequence requires several re-screenings to be fully grasped. Welded Sheep Fence Method Sequence: No specific comments (not applicable method for team).</td>
<td>Connection/Support Materials: We are still unsure of position of wire, inside or outside of posts? How does the wire’s thickness influence overall quality of fence? Includes several novel expressions that are not explained. Cannot see lower part of “stötta”(support) - sequence lacks adequate information. Welded Sheep Fence Method Sequence: Thinks Gripple method insufficient - decided not to follow instruction. Instruction lacks information about electrical wire’s insulation that would prevent it discharging into sheep fence.</td>
<td>Visibility - Knot tying has poor visibility. Visibility/Verifiability/Recordability - Video footage cannot show what is beyond visual obstructions.</td>
<td>Locate and include close-ups that display the wire-tying procedure. Locate and include footage that shows how the “stötta” (support) is formed before it is put into the ground. Add graphics that explain/translate novel expressions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comparability/Verifiability - audiovisual communication is not relatable to users’ preunderstandings about electrified sheep fences.</td>
<td>Briefly, expand on alternative methods for stretching a sheep-fence. Locate and insert footage that communicates isolation-methods.</td>
</tr>
</tbody>
</table>
initially being meant as a sequence that would embody the "how” and “why” communication ideals, as it turns out, have become almost entirely void of action and movement and basically consist of two people talking and pointing, saying things like “do this, do that.” In addition, the many talking points, neatly strung together in a constant flow of verbiage, are at two instances accompanied with on screen text that conveniently fills in some missing information, a clear violation of the modality principle that recommends the use of words as audio narration rather than on-screen text (Clark & Mayer, 2016, pp. 113–127).

Overall, this implies a mismatch between instructional approach and designerly ways to support this approach. However, information richness per se is not the issue here, neither is users’ processing of speech and on-screen text. Instead, we suggest that what the speech and on-screen text communicate—implied actions, problem solving, distances and spatial relations—is the issue (Sweller et al., 2011, pp. 219–233). The point is that this biologically primary information is talked about, never shown, and, therefore, does not effectively provide a means for other instructional techniques that aim to communicate biologically secondary information, that is, instructional video content that requires highly explicit guidance. In other words, the relatively low attention levels of the Isolation and Power sequences is a failure to exploit the inherent human capacity to handle information related to the realm of biologically primary information, by not showing movements and spatial configurations that are imitable and fundamentally recognizable (for humans). The problem is that verifiability most likely is compromised. Moreover, the relatively static camera in both of these video sequences/chapters does not make possible the communication of a (implied) 3D space that most likely would tap into the communication potential of comparability and verifiability. If the camera moves, new space configurations become visible and new comparisons can be made that, in turn, may be more easily verified.

Overall, it seems that the results discussed above ascertain that well-known, laboratory studies, more or less ecological in nature, that indicate that learners’ procedural understanding increases when using the live-action format, may be generalized to real-world learning scenarios. It is a case in point that the sequence associated with the highest attention levels is one of the most unambiguous communication examples of a procedural instruction (the Gallagher knot-tying sequence) in the video. This is consistent with previous research findings with regards to motor skills communication efficacies relatable to the display of hand-movements in live-action videos (Berney & Bétrancourt, 2016; Boucheix & Forestier, 2017; Boucheix & Schneider, 2009; Castro-Alonso et al., 2015; Ganier & de Vries, 2016; Höfﬂer & Leutner, 2007; Lowe & Schnottz, 2008; Marcus et al., 2013; Tversky et al., 2002). It is also important to note that the way the Gallagher knot-tying sequence is presented on YouTube is the result of speciﬁc user feedback that indicated that visibility initially was a great concern. Without visibility, users’ necessary mirror-neuron activation would be compromised (cf. van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009). This particular sequence is also a good example of when recordability is adequately controlled. This sequence displays little, unnecessary/irrelevant information that could have resulted in a cluttered appearance. In part the explanation for this decluttered appearance is the videographer’s choice of framing that is rather close up. This also render high visibility, though it could have been higher if a more high resolution camera (720 p. is lower resolution than HD video) had been used. The results of this particular case study shows that the usership prefers larger screens than smartphone screens. This, we think, is no coincidence as visibility is augmented by high resolution screens that are relatively big.

In brief, then, the primary insight here is that users’ cognitive processing and visual decoding depend on the power of video to show actual human behavior and action. If only telling via the means of text added graphics or audio narration would suffice, this power would go largely unused. A good live- action TC video should therefore include humans and human movement and focus on doing, more so than explaining. This is to acknowledge that purposeful live-action instructional videos use biological primary information as a means to free up cognitive recourses for the really demanding parts of any instruction, which exists in the realm of biological secondary information. In
other words, bottom-up perceptual processes and biologically primary information, make possible, sound socioculturally informed instructional practices.

**Limitations and benefits**

Although the analysis of the cognitive and perceptual design ramifications of the video discussed in this article provide a good starting point for delineating the communication potentials of the live-action format, it does have two limitations, in particular.

First, this case study may be considered atypical, because there are many forms of TC how-to videos and many different kinds of audiences that use such videos for different reasons. This highlights the research problem of wanting to “get inside” the audiences’ heads to figure out how they relate the stated skill acquisition end goal, by the instruction to schemas already held in their long-term memory (Ganier & de Vries, 2016). This also highlights the problem of any media design case study’s capacity to transcend the vagaries of instructional design fashion (cf. Ware, 2013) as well as the challenges associated with possible sources of contingency and non-representation, such as types of work task, visual literacy capacities, and organizational cultures.

Second, the data sample presented in this article is relatively small and does not include unambiguous performance measures that may be considered psychologically or and physiometrically valid. Neither does this study employ a more conventional usability testing method (the way usability experts do usability testing in industry). However, we think these limitations can be overcome through the robustness of this research study’s methods and theoretical perspectives. This, in combination with the appropriateness of analytical tools, research process and data, we think, ensures valid results. These results, in turn, may be analytically generalized by judging the extent to which the findings may be similar (or different) to other instructional assessment studies that involve same kinds of instructional media and similar instructional situations. However, we expect that scholars of visual instructions will elaborate on this study’s overall research approach, including its assessment approach. Studies of greater scope that include other performance and decoding measures will contribute to a greater clarity about live-action communication efficacies. Mobile, in the field-compatible Eye-tracking (ET)-systems could prove useful in this regard. We also expect communication scholars to further refine the use of YouTube analytics that, if at all possible, would diminish the influence of search engine optimization (SEO) related aspects that may factor into the YouTube viewership.

Despite these limitations, we see several benefits, especially with regards to the professional use of the live-action, instructional, video format. Although it is not in the scope of this article to propose a set of design guidelines particularly geared toward instructional live-action videomaking (this, we believe would be somewhat premature), here, we will present a few basic, inter-related, designerly advice.

First and foremost, one good advice is to use the live-action format if the instruction involves people that move. If not, opt for other media formats, such as diagrams. Second, we urge technical communicators involved in live-action videomaking to actively and cautiously manage the application of any designerly additions to the original, recorded, video material. The primary reason for this is that immediacy—what might be labeled yet another key affordance of the live-action video format—may be obstructed by the addition of graphics and audiovisual information that has been “densified.” This might appear as banal advice. Yet we need to remind ourselves of that in the computerized world of audiovisual design, it is often a lot more convenient and quick to “fix” problems in postproduction, rather than engage in the often arduous and costly activity of reshooting instructional video content until it permits effortless decoding on the part of its users. We should also keep in mind that good, recorded video content permits editing that does not affect continuity and, in turn, does not increase cognitive load. It is also a good idea to identify possible element redundancies before applying any design guideline. Design principles’ validity notwithstanding, most live-action, instructional content (that includes humans) already naturally embodies cueing techniques, segmentation techniques,
informal audio narration, and so on, that is, key elements that make up the repertoire of designerly ways to reduce cognitive load and guide attention, just as in good classroom lectures.

Moreover, plentiful artificial design elements in live-action instructional videos may not only cause split-attention and increase cognitive load, but also add unwarranted meaning at the semantic level (cf. LeMarié, Eyrolle, & Cellier, 2008). With regards to so-called lecture videos (e.g., Chen & Wu, 2015; Loh, Tan, & Lim, 2016) this might explain why the implementation of cognitive theory of multimedia learning (CTML) informed design guidelines failed to generate better learning outcomes among students, though the assumption is that they should have. However, there were other positive effects, such as shorter video running times (Lamb, 2015).

So, in summary, then, technical communicators engaged in live-action videomaking would be better advised to regard videography as a design nexus, not just an instructional design template, and thereby stay true to the fundamental tenet of cinematic storytelling: show, don’t tell. And this showing involves movement, in front of the camera, and with the camera (that is controlled movement, not shaky-cam movement). As documentary scholar and educator, Michael Rabiger (2009), reminds us, if only telling would suffice, the “power of cinema to show action and behavior would go unused” (p. 48).

Note

1. As of March 2018, the video can be accessed via https://www.youtube.com/watch?v=vqOnH7QZPHw&t=320s.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Per Erik Eriksson is the program director of Short Story Audiovisual Design at Dalarna University, Sweden. He is a former TV producer and videographer with a special interest in various instructional video/TV formats. He has an MA in Communication from Stanford University (2003) and a PhD (2018) from Mälardalen University, Sweden, in the research field of innovation, design and engineering.

Yvonne Eriksson is professor and vice dean of School of Innovation, Design and Engineering at Mälardalen University, Sweden. She is also the chair of the Information Design Research group and head of MDH Living Lab@IPR. The main interest for her research is communication between individuals and groups in complex organizations. Focus for her research is visual communication and visual management and how the relation between the perceptual and cognitive processes that are involved in interpretation of visuals and build milieus are informed by cultural context and heritage.

ORCID

Per Erik Eriksson http://orcid.org/0000-0002-5921-0227
Yvonne Eriksson http://orcid.org/0000-0001-9057-0063

References


Watson, G., Butterfield, J., Curran, R., & Craig, C. (2010). Do dynamic work instructions provide an advantage over static instructions in a small scale assembly task? Learning and Instruction, 20(1), 84–93. doi:10.1016/j.learninstruc.2009.05.001

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. Learning and Instruction, 22(6), 449–457. doi:10.1016/j.learninstruc.2012.05.004
