

LFI Working Paper No. 6/2016

## VIDEO ELICITATION INTERVIEWS IN ORGANIZATIONAL AND MANAGEMENT RESEARCH: APPLICATION IN A FIELD STUDY

Alexandra Zehe<sup>1</sup>

Ludwig-Fröhler-Institut<sup>2</sup> / Technische Universität München<sup>3</sup>

Frank-Martin Belz<sup>4</sup>

Technische Universität München<sup>5</sup>

---

### Abstract:

Despite growing interest in video-based methods in organizational and management research, the application is rare. In this paper we focus on videos as stimuli in interviews. We compare video elicitation interviews to other forms of interviews, which employ photos as stimuli or which are purely word-based. We suggest five stages of how video elicitation interviews might be applied. Against the background of a field study we share some methodological insights, and discuss the possibilities as well as limitations of video elicitation interviews. We find that organizational and management research could benefit from the inclusion of this method, particularly when exploring sensitive topics, emotions, or identities.

**Keywords:** Interviews, qualitative research, video elicitation

---

---

<sup>1</sup> Mail: zehe@lfi-muenchen.de, Phone: +49 89 515560 83

<sup>2</sup> Ludwig-Fröhler-Institut, Forschungsinstitut im Deutschen Handwerksinstitut (Das Deutsche Handwerksinstitut e.V. wird gefördert vom Bundesministerium für Wirtschaft und Energie auf Grund eines Beschlusses des Deutschen Bundestages sowie von den Wirtschaftsministerien der Bundesländer und vom Deutschen Handwerkskammertag), Max-Joseph-Straße 4, 80333 München, Germany

<sup>3</sup> Technische Universität München, Arcisstraße 21, 80333 München, Germany

<sup>4</sup> E-Mail: frank.belz@tum.de

<sup>5</sup> Technische Universität München, Alte Akademie 14, 85354 Freising, Germany

## 1 Introduction

Despite growing interest in video-based methods in organizational and management research (Bell & Davison, 2013; Slutskey, 2015), the application is rare. Take, for example, the *Academy of Management Journal* (AMJ) as one of the leading management journals. Between 2000 and 2015 a total of 1,168 articles was published in AMJ. Interestingly, only one (!) employed videos as a method of empirical inquiry. In their study on creative group work Harrison and Rouse (2014) use videos as a stimulus for focus group discussions. The rare use of videos in organizational and management research is in sharp contrast to the explosion in the prevalence of videos in modern, contemporary society (Bell & Davison, 2013). With the advent of digital cameras and more recently well-equipped smartphones, taking, processing, editing, and disseminating any kind of visuals became relatively easy, instant and inexpensive (Belk & Kozinets, 2005; Harper, 2005).

Generally, there are two options to employ visuals in organizational and management research, including visual content analysis and visual elicitation (Bell & Davison, 2013). While acknowledging the merits of visual content analysis in organizational and management research (e.g. Benschop & Meihuizen, 2002; Munir & Philips, 2005), we focus on visual elicitation in this paper. More specifically, we describe and discuss video elicitation interviews, which use videos as a stimulus in interviews to elicit opinions and to evoke emotions. We find that the use of videos in combination with interviews holds promise for organizational and management research, which involve emotional or sensitive topics.

The present paper has three main aims. First, it illustrates the relevance and the potential of video elicitation interviews by setting out its key characteristics and differentiating it from other forms of interviews. Second, the paper outlines five stages of how video elicitation interviews might be applied, drawing on the authors' experiences of a field study. Third, some limitations and pitfalls of video elicitation interviews are discussed.

## 2 Key Characteristics of Video Elicitation Interviews

Video elicitation interviews belong to the family of visual elicitation interviews, which employ images as a stimulus in the context of an interview (Harper, 2002; Pauwels, 2015). 'Visual elicitation' is an umbrella term for the use of any kind of image as an interview stimulus, including moving images like videos and still images like photographs (Bell & Davison, 2013). Visual elicitation is a kind of projective technique, which allows individuals to reveal their experiences and meanings by giving them a field (objects, materials) with relatively little structure and cultural patterning (Frank, 1939). Projective techniques originate from the field of psychoanalysis, which assumes that humans engage in conscious, but also unconscious mental processing. The individuals can project upon that plastic field their ways of seeing life, their meanings, significances and feelings (Frank, 1939).

The presentation of images to single or multiple respondents can provide two kinds of information. First, the interview with visual materials can offer the researcher a simple way of obtaining information about whatever is visible in the images (Pauwels, 2015). The visual elicitation interview can provide better and more detailed information on a topic than word-based interviews. When it relates to the past, visual can bring back and sharpen memories, creating richer and more precise information than interviews, which are purely word based. Second, and even more important, the interview with visual materials allows triggering deeper perceptions, values, and emotions of respondents as individuals who are involved in the depicted world (Pauwels, 2015). Carefully chosen images combined with good interview technique can deepen the interview from mere information around the obvious to the meaning of the recorded materials ascribed by the respondents (Pauwels, 2015). Visual elicitation interviews can stimulate memory and evoke emotions in different ways than standard, words-alone interviews (Collier, 1957; Harper, 2002).

According to the form of visual employed as interview stimulus (moving vs. static image), and the source of the visual (researcher vs. respondent) we can differentiate four types of visual elicitation interviews (see table 1). While researcher-generated visuals are more appropriate for theory-driven (or at least theory-inspired) research, respondent-generated visuals are suitable for inductive research approaches (Clark-Ibanez, 2004). In the following we will describe each type, highlighting elicitation studies in organizational and management research (and related fields such as marketing and consumer research).

Source of Visual	Researcher	Respondent
<b>Form of Visual</b>		
<b>Moving</b>	I	II
<b>Static</b>	III	IV

Table 1: Typology of visual elicitation interviews

The first type of visual elicitation interviews relies on moving images provided by the researcher. It refers to short video clips or film sequences, combining images with language and sound. In general, videos can be retrieved from archives or produced by the researchers themselves for the purpose of the study. Videos stimulate different forms of processing and rekindle a range of memories that might otherwise remain latent (Starr & Fernandez, 2007). Short videos or film sequences represent reality in all its intricacy and ambiguity, which offers the possibility to match existing theory and beliefs with lived experiences (Bell, 2008). Such triggers are especially suitable for highly complex settings, where multiple variables influence processes and outcomes, such as interpersonal relationships in an organizational context. Despite the potential, the usage of video as stimulus in interviews remains quite rare in organizational and management research. A notable exception is the paper by Harrison and Rouse

(2014) on the elastic coordination in creative group work. They dealt with the fundamental question of how creative groups coordinate for creative work, allowing individual work and autonomy on the one hand, and group work on the other. They investigated the research question in the context of modern dance groups, which place a premium on creativity (Harrison & Rouse, 2014). In addition to formal, word-based interviews with the individual members they recorded videos on the performance and showed them to the entire group to elicit discussions and reflections on the creative processes. In the related field of marketing and consumer research Sayre (2006) employed video elicitation interviews as well. She conducted a study on the purchase process following a life-changing event (e.g. natural disaster). She produced a realistic video with two actors talking about the feelings and thoughts of fire survivors. In the video elicitation interviews the respondents felt (strongly) connected to the video actors. They tended to bond and identify themselves with the video actors, who seemingly shared the same fate as they did. The respondents developed a kind of “parasocial relationships” (Horton & Wohl, 1956) with the actors in the shown video, which elicited emotions and information on the purchase process following a natural disaster (Sayre, 2006).

The second type of visual elicitation interviews is based on moving images, which are generated by the respondents themselves or in collaboration with the researchers. In her ethnography study on labor Slutskaya (2015) made a collaborative video with working class men doing dirty or undesirable work to overcome suspicion and gain trust. The follow-up interviews with the ‘difficult to research’ groups gained far more information and insights than the first round of interviews, which were elicited with photos provided by the researcher (Slutskaya, 2015). In the related field of marketing and consumer research Starr and Fernandez (2007) developed a method called the ‘Mindcam’, which allows consumers to record shopping experiences from their own perspective. The recorded video are then used as a memory prompt to elicit respondent’s reactions. In their exploratory study the respondents were able to talk in great detail about what they were doing or feeling, which was in part due to the visual nature of the stimulus (Starr & Fernandez, 2007).

The third type of visual elicitation interviews is characterized by still images such as photos, paintings, or public displays (e.g. advertisements, brands and logos) provided by the researcher. Photo elicitation is probably the most common type of visual elicitation interviews (Harper, 2002). It has a long tradition in anthropology (e.g. Collier, 1957; Gates, 1976) and sociology (e.g. Harper, 1984; Steiger, 1995). Collier (1957) introduced the technique of photo elicitation. As part of an anthropological study regarding the relation of the environment on mental health he conducted a (more or less) controlled experiment. The research design involved four informants, whereby two were interviewed with the aid of photographs, while two were interviewed solely with verbal questions. He reports that the photo elicitation interviews provided considerably more concrete information on the structure and processes of the work environment, more specific information on the workers, and more empathetic expressions for certain aspects of industrial work than standard interviews (Collier, 1957). Photographs as static images extend along a continuum between the objective and the subjective (Harper, 2002). At one



extreme there are visual inventories of objects, people, and artefacts. At the other extreme of the continuum photographs portray the dimension of the social, including the own body, the family or other intimate social groups. In between the two extremes are images that depict events of the past, which stimulate the memory of the respondents (Harper, 2002). While photo elicitation interviews have a long tradition in anthropology (e.g. Collier, 1957 & 1967) and sociology (e.g. Harper, 1987 & 2002), they remain rare in organizational and management research (Clarke, 2011). There are some organizational studies, however, employing static images to elicit interviews. Buchanan (2001) documented the process of surgical in-patient care from referral to the eventual discharge by means of photographs, which were presented to five groups of hospital staff at all levels. This led to further amendments to the written account of the process and generated a wider discussion of reengineering possibilities (Buchanan, 2001). In a similar fashion, Ray and Smith (2012) took photos of a business process, followed by photo elicitation interviews with employees and managers about the strategic priorities of the company. However, the photo-elicitation interviews did not unfold as planned due to the selection of photos made. Subsequently, the respondents made suggestions for photos in line with the research project, which led to a kind of hybrid photo production and better empirical results (Ray & Smith, 2012).

The fourth type of visual elicitation interviews uses static images, generated by the respondents themselves. Interestingly, we could identify two empirical studies on organizational change employing drawings as a way to obtain sensitive information. Zuboff (1988) asked employees to draw pictures showing how they felt about their jobs before and after the installation of a new computer system. The drawings elicited follow-up interviews and enabled the respondents to articulate their implicit feelings (Zuboff, 1988). Similarly, Vince and Broussine (1996) used drawings to elicit interviews on emotions and relations underlying organizational change. Warren (2005) asked employees to take pictures of anything in the working environment that mattered to them and that they wanted to talk about. The aim of her study was to explore the impact of the work environment on the feelings and experiences of those who worked there. She argues that the pictures generated by the respondents gave them a louder 'voice' via the accessibility of the method, control of the agenda and the ownership of the resulting image (Warren, 2005). In this context it is to note that the presentation of images potentially changes the relationship between the researcher and the researched. In (more or less) structured interviews the researcher is the 'interrogator' posing questions, while the respondents is the object of interrogation, implying a hierarchical relationship. In visual elicitation interviews the respondents become a knowledgeable informant or even an expert, while the role of the researcher is mainly to steer the conversation and listen to the respondent. It involves a more hierarchical or even equal relationship between the researcher and the researched. This especially holds true for respondent-generated images, which empower the respondents and give them a voice (Pauwels, 2015; Warren, 2005).

## 1 Implementing Video Elicitation Interviews

As already argued, in contrast to purely word-based and conventional interviews, the use of visual stimuli in an interview situation can enhance the generated data regarding its richness and deepness by uncovering implicit and unconscious attitudes, perceptions, and beliefs on part of the respondents. Visuals can also serve as trigger to reveal buried memories and support interviewees to reconstruct the past. Distinguishing moving from still visuals, we especially focus on videos as stimuli in an interview situation. Despite its promise, we identified just a few video elicitation interview studies, and we still lack some guidelines for how to carry out this particular form of interview. Against this backdrop we suggest five steps involved in the planning and implementation of a video elicitation interview study in the context of organizational and management research (see figure 1). We illustrate each step by providing information and sharing insights from an empirical study on the succession in SMEs (Zehe, 2016), employing video elicitation interviews as the main method of inquiry. The study aimed at identifying influence factors on the successor's legitimization as future leader of the business and at examining how family firms manage innovation and change during succession. A multiple case study with ten businesses from the German crafts and trade sector was carried out including video elicitation interviews that were conducted with the predecessor, the successor, and two employees in each firm. The family firms had either accomplished a succession recently or were currently processing one. As family firm succession is a very sensitive, emotional and often conflict-laden topic for all parties involved, questions arose how to gain access to the participants so that they would open up in the interviews and felt up to talk about their personal situation and feelings. As the usage of visuals seemed to be a promising vehicle to dig into even delicate topics, video elicitation interviews were adopted as the most appropriate. Two different video triggers were selected, one shown to 16 interviewees, the other one to 36 respondents.

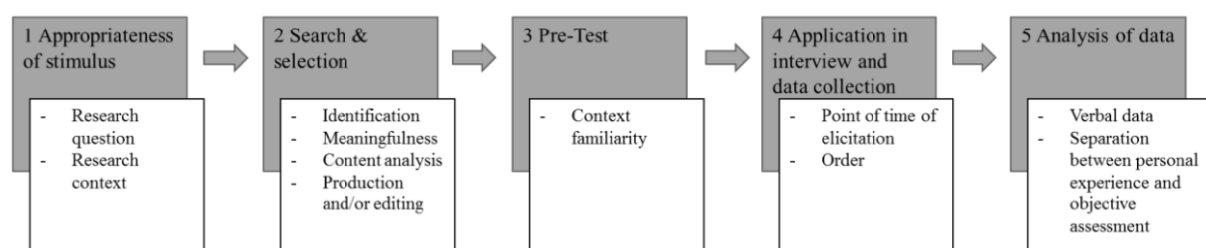


Figure 1: The five stages of implementing visual elicitation interviews

### Step 1: Appropriateness of Stimulus

In the beginning the type and the source of the stimulus has to be decided (cf. table 1). This decision depends mainly on the research question and the research context. For instance, researcher generated or

archival visuals might be useful to trigger memories, to reconstruct past events (Collier, 1957) or to analyze processes (Buchanan, 2001). In contrast, respondent generated or participatory produced visuals support participants to emphasize what they believe is important by giving them a “voice” (Pauwels, 2015) and by creating trust between researcher and respondent (Ray & Smith, 2012; Slutskaya, 2015). Also, respondent generated visuals are limited to the current state of lived reality, because the generated visuals are recorded in real time and cannot restore the past. Comparing video and photography, Collier and Collier (1967) emphasize the benefit of video in contrast to photo referring to its capacity to present the quality, nature, and composition of social behavior and human relationships that cannot be captured in the same deepness with frozen photographs. The emotional and communicational aspects of interpersonal relations manifest themselves in the characteristics of the video such as motion, sound, language, and interaction. Thus, if the research questions aims at capturing emotional and sensitive topics or feelings, or strives at evaluating interpersonal relations in complex settings, still photographs might not be appropriate to depict the fine nuances and subtleness of human behavior, because they contain refinements that provide a great deal of interpretative leeway (Collier & Collier, 1967).

The Succession Study used researcher generated videos that were created by editing specific scenes out of existing movies for several reasons. First of all, succession in family firms is often a highly emotional and conflict-riddled situation for all parties involved. Catching individual feelings and attitudes toward multi-faceted topics in a complex environment by applying photographs would not have been expedient. Furthermore, one research question aimed at investigating a specific, interpersonal situation between predecessor and successor that has happened in the past, which is why researcher generated material was used. Instructing the respondents to produce their own video clip with regard to this past situation would have been inappropriate to elicit revealing information as it might have resulted in a simple renarration that lacks, in the worst case, important aspects. Lastly, drawing on pre-existing film material seemed to be a rather economically priced option than producing an own video clip. We only had to find and select appropriate movies that included exactly the aspects from the research questions we were interested in.

## **Step 2: Search for and Selection of Stimulus**

In the next step, the stimulus itself has to be chosen. In case of researcher generated videos, strong emphasis has to be laid on the accordance between stimulus content and research context and questions, respectively. Although referring to still visuals, Ray and Smith (2012) stress the importance of selecting “meaningful” photographs (Ray & Smith, 2012: 310), otherwise interviewees will not be able to respond to the photos and the trigger fails. According to Collier and Collier (1967), by considering the complete emotional and evocative content of a photograph, powerful, expressive, and provocative ones are the most effective in terms of an efficient projective response. These aspects are also applicable and valid for the selection of videos. Furthermore, it is important that respondents can identify with the videos presented to a great extent. Sayre (2006), for instance, uses self-produced video clips with professional

actors in order to gain access to natural disaster victims. If the actors would not have been rated as authentically by the interviewees, the development of “parasocial relationships” (see Horton & Wohl, 1956)—a strong bond between interviewee and actor—would have been unsuccessful and the trigger video useless. Potential triggers can be identified by carefully searching several media canals, for example in web-based media libraries. Due to the growing presence and importance of videos in social and digital life, access to this type of media has become far easier and cheaper. After searching for movies, films, or videos, based on keywords that deal with the research context and identifying potential ones, the selected stimuli have to be analyzed in deep regarding their content, meaning, and hidden information. If several possible triggers have been chosen, they should be compared by weighing up the advantages, disadvantages and anticipated reactions, completed by the selection of the most appropriate one(s)—always with regard to the research questions.

Within the Succession Study, we first ran searches in the internet to identify a number of movies from the past years that dealt with family businesses in general. We then scanned the search results and eliminated those movies where succession did not play a predominant role. The remaining and promising movies were afterwards bought, watched closely and examined, whether the specific aspects we aimed at were included. Highly potential scenes were noted down and watched again. This was followed by a detailed microanalysis of each possible scene including its storyline, discourse, its symbolic and hidden meaning.

After comparing the identified film scenes with the research questions thoroughly, we chose two different movies, selected one scene from each film that seemed to fit best and edited the stimuli out of the movies. The first movie was the Danish drama „A Family“ (Jørgensen, Wiedemann, & Fischer Christensen, 2010) that focuses on the unsolved succession problem in a traditional Danish family bakery and depicts the interaction between successor and predecessor. According to Bordwell and Thompson (2004), the plot of the movie and its main events contain the more obvious and referential meaning of a film, which is why the basic story line of the trigger is provided in the following video vignette 1.

#### **Video Vignette 1: “A Family” (2010)**

The opening montage emphasizes the longstanding tradition of the family bakery by using single text modules alternating with monochrome images accompanied by a sleepy, melancholic background music. It provides a wealth of information about the family’s back story. Values and ideals such as familiness, quality of the pastries, and long experience in the prestigious business come to mind naturally.

The film scene itself shows a conversation between the predecessor, who is fatally ill, and his daughter, the potential successor, who is not aware of the father’s plans and wishes. The predominant silence and absence of music in the scene is quite a contrast to the opening trailer. The scene starts with the

predecessor being alone in the shop floor, selecting logs of wood for the baker's oven, lighting the fire, sitting down in front of it, and staring into the oven for while. There is long silence, only the wood is cracking. Then his daughter arrives, sits down beside him, without having any clue about what the predecessor wants to talk about as he had pleased her to come. The predecessor starts alleging reasons why he sees her and nobody else stepping into the business. She is the only capable successor he can imagine in his private sphere, listing all shortcomings of other potential heirs and employees. The conflict becomes evident as the daughter reveals her personal life plans, while the predecessor reproaches her: "You said you would always be there for me." Also, he reminds her of the family tradition that runs the risk of getting lost. The scene closes with the predecessor embracing his daughter, kissing her on the cheeks and whispering her first and last name, which reveals his pride of the family name and shall obviously remind her to be aware of the name she bears.

The second trigger stemmed from the movie "The Buddenbrooks" (Abich, Krause, & Weidenmann, 1959), which pictures the rise and fall of a German merchant family based on Thomas Mann's famous novel. In the stimulus, focus is laid on the interaction between the successor and an employee. Likewise, video vignette 2 gives an overview of the story line of the stimulus. An image sequence of both videos is provided in Appendix 1 and Appendix 2.

#### **Video Vignette 2: "The Buddenbrooks" (1959)**

In the beginning of the video, a freeze image in form of a short text is presented, which introduces the circumstances of the succession to the interviewee: "the predecessor has died, which is why the successor steps in; the following clip refers to a situation where the new successor talks to an executive employee".

The actual clip starts with the successor talking to the executive employee while hurrying into his office. The conversation is about the successor's plans regarding innovations to continue the business: for instance, investments in a merchant fleet are necessary to remain competitive. At first, the employee is quite surprised and worried that such steps might be too far-reaching and too risky to handle financially. But the successor makes reasonable points that persuade the employee finally. The scene ends with the successor opening the window and saying "you have to move with the times or the times move with you".

A surface impression of the analysis hints that both clips are quite different from each other if not even constitute polar cases. For instance, the video vignette taken from the movie "The Buddenbrooks" stems from 1959 and is all black-and-white. Sound and voices are very fast and harsh, which makes a hectic impression on the watching person. In contrast, the vignette from "A Family" is a relatively modern

movie, which finds expression in details such as camera work, color film, and resolution. The “A Family” video clip was a very long stimulus with about five minute’s length, while “The Buddenbrooks” video was shorter with only one minute 40 seconds. Both clips were split into an opening montage, which contained some background information about the business’s situation, and a main part that presented the actual topic.

Both clips aimed at stimulating and eliciting a specific topic: “A Family” was presented in order to find out when the first conversation about the possible entry of the successor took place and how the decision was eventually made. Still, the clip covered in many aspects the whole spectrum of a succession. In contrast, the vignette from “The Buddenbrooks” dealt with the successor’s innovative behavior and aimed explicitly at analyzing the respondents’ view on innovation and change during succession processes.

On a deeper level of meaning, “A Family” takes up the successor’s struggle and conflict with taking over the prestigious family business. She is caught in a dilemma between serving her father’s expectations to continue the business and saving the family tradition or following her own dreams. Thus, the scene gives an idea how conflict-laden, tragic and tense the family and business situation is—without solving the conflict in the end. The actors represent also “equal” partners in their double role as father/predecessor and daughter/successor. Moreover, the scene contains a lot of symbols, such as the fire, the ashes, and the baking oven.

In contrast to that, “The Buddenbrooks” video describes on a level of explicit and implicit meaning (Bordwell & Thompson, 2004) the situation every successor faces when entering a business: Successors are often caught in a dilemma between preserving existing structures and implementing new ideas and innovations. Furthermore, in this scene, the employee and successor represent a rather hierarchical superior-subordinate relationship. In the end of the scene, both agree upon necessary changes in the future—thus, the initial problem is solved and the scene ends rather harmonic. The main symbols used are ships and an opened window.

### **Step 3: Pre-Test**

When video triggers are researcher generated or stem from archives, showing the selected videos in a pre-test to probands secures that the desired reactions are evoked effectively. If the pre-test participants do not fulfill the preconditions that were set for the later interviewees or are not familiar with the topic, they should at least be requested to put themselves in the interviewees’ place. This allows to verify, whether the interviewees refer in their interpretation of the shown situation to the same context and facts comprised in the research question, and whether information is revealed the researcher actually wanted to trigger. If the pre-test evinces that the trigger does not create oral fluency or that responses lead in the wrong direction, the trigger should be edited or rethought. At the extreme, it is reasonable to consider returning to the selection stage and choosing another one.

In case of the Succession Study, an internal pre-test with two fellow researchers was conducted. One of them had a family firm background and could therefore better empathize with the actors in the video stimulus. After the promising pre-test, it was decided to use both triggers for the interviews within the first case. The four interviewees from the first family firm served therefore as further pre-test. Reactions observed were satisfying and the videos triggered a flow of words providing interesting and new insights aligned with the research questions. Therefore it was decided to keep both triggers for further interviews without editing or changing them.

#### **Step 4: Application in interview and data collection**

After a successful pre-test, the researcher generated triggers can be applied in the interview. First, the researcher has to decide, *when* the trigger will be presented and in case of multiple stimuli, in which *order*. Also, the researcher should make sure that the technical equipment works smoothly, otherwise the course of the interview can be disturbed considerably. After presenting the video, the researcher asks the interviewee for an interpretation of the watched film by posing a reasonably open question.

In the Succession Study, one interviewer conducted single focused interviews with predecessors, successors, and employees in each firm on-site between September 2013 and March 2014. The order of the interviews was arranged in situ and determined by the owner-managers. Thereby, the participants did not know beforehand that videos were used in the interviews. The duration of the interviews ranged between 25 minutes and 2 ½ hours, whereas the interviews with employees were mostly shorter than with the owner-managers. In general, the video vignettes were presented on an iPad to the respondents. While the video vignette “A Family” was shown quite in the beginning of the interviews to a total of 16 participants, the film excerpt “The Buddenbrooks” was presented rather in the end of the interviews to a total of 36 interviewees. The reason for the difference in the number of elicitation interviews is that the video vignette “A Family” was not shown to the employees, because they have not experienced the depicted situation personally, which is why the video would not have triggered any promising information due to a lack of personal identification. However, with the “Buddenbrooks” trigger, every respondent could identify him-/herself.

The first question after presenting the video vignette was in a very unstructured and open manner, as for example: “What do you think about the film?” or “What comes to your mind?” (see Merton & Kendall, 1946). By asking open and unstructured questions on purpose, the researcher gives the respondents the possibility to react to any aspect of the presented stimulus without pushing them into one specific direction. It was clear that the respondents would not react to the videos similarly as the films held different meanings for each of them. When the interviewees’ assessments and responses to the film situation were surprising or unexpected compared to the previous objective film analysis (cf. step 2), the researcher enquired into the matter more precisely.

### **Step 5: Analysis of data**

The data analysis of researcher-produced video elicitation interviews relies exclusively on the interview-generated verbal data. The data material provides rich, usable, and new indications in the ideal case. Also, the data will not only refer to the referential, obvious meaning of the watched film, but also to the explicit, implicit and symptomatic meaning as Bordwell and Thompson (2004) name the deeper and hidden meaning of movies. Thus, the interpretation of the trigger and its effects might differ between the respondents dependent on their personal situation and experiences. Also, it does not remain an objective evaluation, but is more: The hidden symbols and sayings in the video trigger, the identified and depicted social interrelations with all its power structures, emotions, and its subtleness are mixed with what the respondents have underwent so far as they project their experiences onto the watched movie. The verbal data can therefore be rich and colorful, and contains a lot of information and indications about their view of life, their attitudes, and values.

Although the richness of the data is one of the most important advantages all elicitation interviews hold, it leads to complications when it comes to data analysis. The separation between the objective assessment of the trigger and the subjective revelation of the respondents' own attitude becomes difficult. Hence, the question is: Is only the respondent's elicited personal experience relevant for data analysis or does the apparently objective evaluation of the trigger—as, for instance, the immediate response after the researcher's open question—contain also important information that might be relevant for later theory development? Which data fragments are selected and how shall these single quotations be sorted into the coding scheme and the paramount categories?

We did not anticipate those problems in the Succession Study and faced them for the first time, when it came to the analysis of the first interviews. After careful consideration, we decided to code the assessments of the trigger videos as well. We did not distinguish between the interviewee's personal experience in a similar situation as depicted in the video and the evaluation of the trigger. Reasons therefore were that we recognized that even a supposed objective evaluation of a trigger contains a lot of subjective information. Thus, from our experience we conducted that every evaluation contains subjective indications and allows inferences, which is why the interpretation of the trigger as well as the respondents' individual experiences were coded the same way.



## 2 Discussion

Starting with some empirical reflections about the application of researcher generated videos in the Succession Study, we want to elaborate in the following paragraphs how the triggers worked, which effects they had and what problems we faced during implementation. This is followed by explicating some methodological reflections regarding this interview method.

### Empirical reflections

As mentioned before, both applied video triggers constitute contrarian stimuli in a certain way. From the resulting, various effects and reactions in the interviews, it was deduced that these differences were causative. We therefore abstracted these distinctions and identified several relevant aspects, which should be taken into consideration when selecting and applying researcher generated videos as visual stimuli.

***Purpose of the trigger: anchor or arrow?*** We observed substantial differences in how the respondents reacted to the two triggers, which is why we created two distinct types of stimuli videos. In the Succession Study, the “A Family” video was termed as *anchor type*. This trigger type provides the interviewees with broad insights into the topic as it alludes and names many different aspects of the subject. It enables interviewees to get back and refer to the film scene during the whole interview (“as you have seen in the movie”) when they want to bring new aspects into play. Another advantage the anchor type holds is that the researcher can easier address new topics as they have probably been mentioned in the respondents’ interpretation of the movie before. Thus, barriers for the researcher to address new subjects are lower and respondents are more willing to open up, because researcher and respondent have already talked about the topic—even if marginally in another context.

In contrast, we termed “The Buddenbrooks” video as *arrow type*, as substantial distinctions were recognized in opposition to the anchor type. As this trigger type focuses on one specific aspect, its application is advantageous if the researcher wants to elicit very accurate responses regarding a specific topic. On the other side, a well-directed stimulus entails the risk of obviousness as respondents might more easily discover the intention behind the scene. In general, anchor type triggers should rather be presented in the beginning of the interview, while arrow type videos can be shown at any time.

***Degree of conflict and disharmony.*** Presenting a video, which contains a certain degree of conflict and disharmony is important as it enables the later discussion of the actors’ produced arguments or at least stimulates the interpretation of the trigger situation on part of the respondents. A trigger without controversies and a saturated level of harmony is inappropriate for eliciting the respondents’ own opinion. In both videos in the Succession Study, a certain degree of conflict was present, which is why the actors’ single arguments could be discussed with the interviewees regarding their substantiation and weaknesses.

Furthermore, these conflicts should remain unsolved and incomplete—closure should be avoided. This open-endedness stimulates the later discussion and gives room for dialogue so that the interviewees can explore ideas and solutions on their own. In the Succession Study, responses to the “A Family” video, where the conflict was not solved, were more detailed and elaborate, while the “Buddenbrooks” video that ended in harmony produced rather short responses. The interviewees evaluated the actors’ behavior rather poorly and a solution different from the proposed one in the trigger was avoided. Thus, in this case, segueing from the video trigger to the respondents’ own in-firm context was initiated by the researcher rather fast and resulted in considerable difficulties to elicit information.

***Expected degree of identification.*** We found that the presented video must contain a situation the respondents can easily identify with. At best, they have experienced the concrete situation themselves, although this is not something the researcher can anticipate. But this enables the respondents to put themselves into the actors’ place, which in turn stimulates memories of their personal experience in a situation like this or at least enables them to judge how they would have reacted. This effect is also enhanced if the chosen stimulus contains a plausible, realistic, and credible setting (Buchanan & Huczynski, 2004). Sayre (2006) emphasizes the importance of establishing such unilateral, “parasocial relationships”—the term stems from psychology (Horton & Wohl, 1956)—where watchers identify and associate with the actors, feeling acquainted and strongly connected with each other.

We observed that the higher the degree of identification with the actors was, the more the respondents were willing to give insights into their personal experiences and tended to compare their problems with the presented film situation. Although not every successor and predecessor had experienced an identical situation as presented in the “A Family” trigger, both interviewees could at least tell the researcher how the first conversation between the incumbent and the heir took place and how they morally judge the shown situation in the trigger. In contrast, “The Buddenbrooks” did not work equally well, because not every employee had experienced such a conversation with his/her boss similar to the presented one in the video. At least, the interviewees gave insights how decisions on the implementation of innovations in the business were managed.

The requirements regarding the degree of identification confront the researcher with a dilemma: on the one side, a video trigger has to be chosen, which enables the highest degree of identification. This requirement calls for selecting a specific situation the individual has experienced. On the other side, the likelihood for such an identification on part of most of the interviewees is increased the broader the presented stimulus situation is, because the researcher cannot anticipate and take the personal background of each respondent into account.

### **Methodological reflections**

Applying researcher generated video triggers allowed us to gain deep insights into the respondent’s personal experience and way of seeing, which words-alone based interviews would not have allowed.

Nevertheless, the researcher selected the situations depicted in the trigger and not the respondent, which means that the researcher did not give the participant the freedom to choose a situation on his/her own that he/she finds the most important regarding the research topic. In this respect, the interviewees were possibly to a large extent driven by the researcher. Königstorfer and Groeppel-Klein (2010) made a similar observation in their study applying photo elicitation interviews where they concede that participants were not exclusively “autodriven” (Königstorfer & Groeppel-Klein, 2010: 405).

In general, it is often argued that the use of visuals in an interview situation mitigates the hierarchical relationship between researcher and respondent in favor of the participant, who becomes a “knowledgeable informant” (Pauwels, 2015: 98) rather than an object of study. We suggest that a more detailed separation has to be made dependent on the source of the visual. While researcher generated visuals bias the information eliciting process in so far as a trigger situation is predefined by the researcher, respondent generated visuals give a greater scope of influence to the participant due to the participative, involving manner of the trigger production. The results obtained by the respondents are out of control for the researcher, who forfeits therefore some of his hierarchical power in guiding the interviewing process. Here, as producer of the trigger the participants highly influence and determine what is spoken in the interview situation. In contrast, researcher generated stimuli disable respondents to select what they believe is meaningful and important. Here, the researcher has enormous influence on directing and channeling the interview situation through the stimulus selection he/she made regarding the research question and context.

This leads to the assumption that researcher generated stimuli are more appropriate for deductive, theory-driven approaches, while respondent generated triggers are rather suitable for inductive, open approaches, as for instance, such studies using grounded theory. Also, video triggers equivalent to the anchor type supposed in this paper provide an easy start into the interview and a broad basis for discussion. These are ideal for open and unstructured interviews that follow an inductive or abductive approach, because they provide reasonable scope for the respondent to react to any aspect the interviewee finds important. In contrast, arrow video types are more appropriate for structured interview situations, where the researcher exactly knows where to aim at. Thus, this goes along with a more deductive procedure.

### **3 Limitations and Pitfalls**

Naturally, every type of visual elicitation interview, respondent as well as researcher generated, photo as well as video, contains its own limitations and pitfalls. In the remainder of this sections, we want to present some pitfalls we faced during the application and data analysis in order to prevent other researchers from the obstacles we met.

**Hierarchy of human relationships.** Every human relationship is characterized by power and influence between the involved people, one exerting more or less power over the other person. This becomes obvious in top-down hierarchical relationships, especially in organizational context. Thus, when selecting the trigger, the researcher should consider the target respondent and his/her direction of assessment carefully—is an upward or downward evaluation requested? For instance, no problem arises, if the respondent is in the superior position and is asked to interpret a stimulus situation that depicts a boss-employee relationship. The reverse setting can be more complicated: Here, a subordinate interviewee is asked to interpret or judge a shown trigger situation that depicts, for instance, an argument between a manager in a superior position and a subordinate employee. Such a situation entails the risk of creating a social desirability bias as the interviewee in the subordinate position might not dare to judge the stimulus scene honestly, as it was the case in “The Buddenbrooks” video in the Succession Study.

But although equally hierarchical relationships presented in video triggers might be easier to handle than top-down relationships, they cannot be avoided completely as the desired reactions the researcher is interested in sometimes allude to this top-down hierarchy in particular. Also, one should keep in mind that videos highly facilitate the depiction of these hierarchies, while still photographs rather fail to represent the fine nuances of power, influence, and emotions in interpersonal relationships. Still, triggers including upward assessments where interviewees, for instance, have to evaluate or even criticize a superior person should be considered carefully and applied cautiously.

**Symbolic content.** Visible, obvious and easily interpretable symbols should be used, as for instance the fire in the film “A Family”. Zaltmann (1997) also emphasizes the importance of being sensitive to metaphors as they are “the engine of imagination” (Zaltmann, 1997: 425). Methods that support metaphoric thinking can elicit hidden knowledge and significantly increase the richness of data in contrast to words-alone approaches (Zaltmann, 1997).

In the “A Family” video, fire, flames, and ashes were often-depicted symbols. Some interviewees referred to those symbols by taking up German sayings or phrases that contain verbal allusions (for instance, one interviewee said: “This predecessor does not pass on fire and passion, but only the ashes”). Here, the symbols were used and packed into metaphors. The baking oven scene is very long, slow and silent and hence, fire becomes a very predominant symbol in the movie. Some interviewees became very nervous and uncomfortable. Thus, as some symbols and the way they are staged can be perceived as very strong, “intense” stimulus scenes as the referred oven scene cannot be recommended in the beginning of the interview. Also, symbols that are not visible but only mentioned, as for instance the ship in the “Buddenbrooks” trigger, are rarely noticed by the respondents and did not show any effect in the later verbal interview data. In the Succession Study, we were surprised by the power the symbols unfold within the participants and the fierceness and intensity the interviewees responded to them. Thus, trigger videos should be carefully examined regarding their symbolic content and hidden meaning before the application in the interview as they will definitely emerge again in the subsequent verbal data.

**Verbal content and sayings.** We experienced that interviewees willingly picked up sayings or repeated phrases, which were used by the actors in the video scene. This was striking in case of the “Buddenbrooks” trigger, where many interviewees recapitulated or complemented sentences, as for instance: “We have to move with the times, otherwise the times move with you.” Hence, the dialogue in the trigger scene should be carefully examined regarding sayings and phrases that provide scope for being recapped by the interviewees. On the one side, it facilitates reactions and supports interviewees that might be more inhibited to divulge. On the other side, it renders possible generalizations and does not force the interviewees to make their own concrete and meaningful statements.

**Length of the scene.** The length of the scene depends generally on the purpose of the trigger. Anchor scenes can be longer as they need to provide much information; in contrast, arrow scenes should be short and come to the point. Still, the longer the video trigger is and the more information it contains, the higher is the risk that respondents are overflowed with too many details they cannot keep in mind and respond to in their later interpretation. On the other side, weak film content is liable to trigger and elicit less information. Thus, the selected scene should be a balanced weighing up of necessary content and required length.

**Demands on respondents.** At least, researchers applying visual elicitation interviews should consider the video content carefully. Because the respondents do not have much time for interpreting the video and do not know the whole story behind, all too demanding videos including complex issues might miss the target as they will not trigger anything if the discourse is not understood. For instance, some employee-interviewees in the Succession Study had difficulties with understanding the storyline of “The Buddenbrooks” trigger in particular, which lead to a misinterpretation of the video. In contrast, the predecessors and successors did not have any problems at all with both triggers. This might show that the “Buddenbrooks” video was not selected appropriately and that the level of employment and education of the target group should also be kept in mind when selecting a potential stimulus.

#### **4 Conclusion**

Visuals have raised considerable interest in many research fields due to their unconventional application in the field and the generation of promising, interesting, and well-grounded results. Visual elicitation interviews enable a new order of dialog between researchers and participants, which is more based on interaction and discussion. Researchers applying visual elicitation interviews hope for richer, truer and less biased data from the respondents, as images are supposed to stimulate other parts of the human brain than verbal or written data and evoke deeper elements of the human consciousness (Harper, 2002).

By sorting existing studies from organizational research into a two-by-two matrix dependent on the form of the visual and its source, we showed that organizational researcher lately use visuals in different forms

more frequently when studying organizations. Still, the usage of static images could be identified as predominant, while the application of moving visuals remains rare.

Based on the scheme of four different types of visual elicitation interviews, the aim of this paper was to give guidelines when which method is more appropriate, dependent on the research question and context. We then specifically picked one specific method (type I) and established a practical hands-on guide for researchers intending to apply video elicitation interviews. The single steps were thereby deduced from a study with family firms from the German crafts sector currently facing succession, where researcher generated video elicitation interviews were conducted.

In step 1, it has to be decided on the appropriateness of the stimulus, in other words, which stimulus is the most appropriate one in terms of research question and context. Here, it is helpful to return to the posed two-by-two matrix, as it provides a systematic scheme, from where the researcher can choose between four different types of visual elicitation interviews. We distinguish between researcher- and respondent generated visuals, which presents the source of the visual, and still and moving images—the form of the visual. Due to the deferral of hierarchical power to the researcher to the disadvantage of the participant in case of researcher generated visuals, this method is especially suited for theory-driven, deductive approaches. Here, the participant is rather driven by the researcher, who knows beforehand where he/she exactly aims at. Regarding the aim of the study, the method is apt for triggering memories, for reconstructing past events, and for analyzing processes. In contrast, respondent generated stimuli are rather suitable for inductive, open approaches, as the interviewees determine the focus to a large extent by selecting important aspects on their own. This is why in this case power is more deferred to the participant, who takes the role of the informant and expert rather than the object of study. Here, the researcher should be aware that participatory methods are predominantly limited to present the current state of lived reality and always have documentary traits.

Comparing still and moving visuals, the latter is more appropriate when it comes to portraying complex human relationships and settings, where still images would not have the explanatory depth conveying quality, nature, and composition of social behavior. A frozen image is not suited to represent the subtlety and shades of human behavior due to its lack of movements that indicate gestures and interactions, and its lack of voice and sound that cast light on dialogue, context, and communication.

After the determination of the source and form of the visual, step 2 comprises the search and actual selection of the stimulus. In a pre-test with respondents familiar with the research topic and context, the chosen or edited video triggers should be tested and verified. It is important that the stimulus evokes the desired responses and does not lead into directions remote from the research question. After a successful pre-test, the video trigger can be employed in the real interview situation where verbal data is collected. In case of researcher generated video elicitation interviews, the generated and recorded verbal data is relevant for later data analysis. Here, in the last step, we tried to sensitize the researcher to the importance

of coding the “objective” assessment of the trigger situation as well as the revelation of his/her personal experience in a situation similar to the trigger video.

Our empirical reflections were mainly based on the comparison of both applied video triggers in the Succession Study, as these stimuli differed from each other in fundamental aspects and constituted polar cases, which finally resulted in various effects from the respondents. We therefore identified two different types of video stimuli—the anchor and arrow type. While the first one offers a very broad situation giving much space for discussion, the arrow type is often shorter and focusses on a very specific aspect. The researcher should also select a video trigger that contains a certain degree of conflict and disharmony, which is not resolved and thus engages respondents to find a solution. He/she should also choose a situation where it is likely that the respondents can identify with, which is the case when they have probably experienced such a situation on their own. Otherwise, responses to the shown situation might be difficult due to the lack of a ‘parasocial’ relationship between participant and actors (Horton & Wohl, 1956). Furthermore, the depicted hierarchy between the actors that shall be evaluated by the respondents, the symbolic content, included sayings and verbal phrases, the length of the scene, and the demands on the target group should be carefully considered and weighed.

From our experience, we argue that video elicitation interviews as semi-projective method are most suitable to reveal values, attitudes, and feelings, the respondents consciously know or have reflected upon. An assessment of the trigger situation and its justification relies to a large degree on the personal situation and experiences the participant has gone through. Thus, every individual will give a unique appraisal of the shown stimulus situation. These experiences as origin of the subjective evaluation of the trigger should be discovered and further discussed in the interview.

Keeping all this aspects in mind, we experienced in the Succession Study that video elicitation interviews were an efficient and helpful vehicle that supported us in gaining fresh and unforeseen insights into the research topic. The major contributions of this paper are that we provide recommendations for researchers who want to apply visual elicitation interviews and are not sure, which method might be the most appropriate. We also equip them with a practical guide for the conduction of researcher generated video elicitation interviews and give a comprehensive overview of pitfalls and limitations of the method in order to alert researchers to possible drawbacks. This contributes to the setting of standards and to the enhancement of the reliability and validity of the method, which could increase the application of the method in organizational research in the future.

## References

- Abich, H., Krause, E. (Producers), & Weidenmann, A. (Director) (1959). *The Buddenbrooks* (Film). Germany: Studiocanal.
- Belk, R.W., & Kozinets, R.V. (2005). Videography in marketing and consumer research. *Qualitative Market Research: An International Journal*, 8 (2), 128–141.
- Bell, E., & Davison, J. (2013). Visual Management Studies: Empirical and Theoretical Approaches. *International Journal of Management Reviews*, 15 (2), 167–184.
- Bell, E. (2008). *Reading Management and Organization in Film*. London, UK: Palgrave MacMillan.
- Benschop, Y., & Meihuizen, H. E. (2002). Keeping up gendered appearances: representations of gender in financial annual reports. *Accounting, Organizations, and Society*, 27, 611–636.
- Bordwell, D., & Thompson, K. (2004). *Film Art: An Introduction*. New York, NY: McGraw-Hill.
- Buchanan, D. (2001). The Role of Photography in Organization Research. *Journal of Management Inquiry*, 10 (2), 151-164.
- Buchanan, D. & Huczynski, A. (2004). Images of Influence. 12 Angry Men and Thirteen Days. *Journal of Management Inquiry*, 13 (4), 312–323.
- Clarke, J. (2011). Revitalizing Entrepreneurship: How Visual Symbols are Used in Entrepreneurial Performances. *Journal of Management Studies*, 48: 1365–1391.
- Clark-Ibáñez, M. (2004). Framing the social world with photo-elicitation interviews. *American Behavioral Scientist*, 47 (12), 1507–1527.
- Collier, J. Jr. & Collier, M. (1967). *Visual Anthropology. Photography as a Research Method*. New York, NY: Holt, Rinehart & Winston.
- Collier, J. Jr. (1957). Photography in Anthropology: A Report on Two Experiments. *American Anthropologist, New Series*, 59 (5), 843–859.
- Frank, L. K. (1939). Projective methods for the study of personality. *Journal of Psychology*, 8, 389–413.
- Friedenberg, J. D. & Silverman, G. W. (2015). *Cognitive Science: An Introduction to the Study of Mind* (3rd ed.). London, UK: Sage.
- Gates, M. (1976). Measuring Peasant Attitudes to Modernization: A Projective Method. *Current Anthropology*, 17 (4), 641–665.
- Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2013). *Cognitive Neuroscience: The Biology of Mind* (4th ed.). New York (NY): Norton & Company.



- Harper, D. (1984). Meaning and Work: A Study in Photo Elicitation. *International Journal of Visual Sociology*, 2 (1), 20–43.
- Harper, D. (1987). *Working Knowledge: Skill and Community in a Small Shop*. Chicago, IL: University of Chicago Press.
- Harper, D. (2002). Talking about pictures: a case for photo elicitation. *Visual Studies*, 17 (1), 13–26.
- Harper, D. (2005). What's new visually? In N. K. Denzin & Y. S. Lincoln (Eds.) *Handbook of Qualitative Research* (2<sup>nd</sup> ed., pp. 747–762). Thousands Oaks, CA: Sage.
- Harrison, S. H., & Rouse, E. D. (2014). Let's dance! Elastic coordination in creative group work: a qualitative study of modern dancers. *Academy of Management Journal*, 57 (5), 1256–1283.
- Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction; observations on intimacy at a distance. *Psychiatry*, 19 (3), 215–229.
- Jørgensen, S. G., Wiedemann, V. (Producers), & Fischer Christensen, P. (Director) (2010). *En familie (Film)*. Denmark: Film i Väst & Zentropa Entertainments.
- Merton, R. K., & Kendall, P. L. (1946). The focused interview. *American Journal of Sociology*, 51 (6), 541–557.
- Munir, K. A. & Phillips, N. (2005). The Birth of the 'Kodak Moment': Institutional Entrepreneurship and the Adoption of New Technologies. *Organization Studies*, 26 (11), 1665–1687.
- Pauwels, L. (2015). 'Participatory' visual research revisited: A critical-constructive assessment of epistemological, methodological and social activist tenets. *Ethnography*, 16 (1), 95–117.
- Ray, J. L., & Smith, A. D. (2012). Using Photographs to Research Organizations: Evidence, Considerations, and Application in a Field Study. *Organizational Research Methods*, 15 (2), 288–315.
- Sayre, S. (2006). Using video-elicitation to research sensitive topics: understanding the purchase process following natural disaster. In R. Belk (ed.) *Handbook of Qualitative Research Methods in Marketing* (pp. 230–243). Cheltenham, UK: Edward Elgar.
- Slutskaya, N. (2015). *Better Together: Examining the Role of Collaborative Ethnographic Documentary in Organizational Research*. Working Paper, Brunel Business School, Brunel University, London, UK.
- Starr, R. G., & Fernandez, K. V. (2007). The Mindcam methodology: perceiving through the native's eye. *Qualitative Market Research: An International Journal*, 10 (2), 168–182.
- Steiger, R. (1995). First children and Family Dynamics. *Visual Sociology*, 10, 28–49.

- Vince, R., & Broussine, M. (1996). Paradox, Defense and Attachment: Accessing and Working with Emotions and Relations Underlying Organizational Change. *Organization Studies*, 17 (1), 1–21.
- Warren, S. (2005). Photography and voice in critical qualitative management research. *Accounting, Auditing & Accountability Journal*, 18 (6), 861–882.
- Zaltman, G. (1997). Rethinking Market Research: Putting People Back In. *Journal of Marketing Research*, 34 (4), 424-437.
- Zehe, A. K. (2016). *Legitimization of the successor in family SMEs: Developing a theoretical framework* (Unpublished Working Paper). Technische Universität München.
- Zuboff, S. (1988). *In the age of the smart machine*. New York, NY: Basic Books.