# COMPLEX DESIGN DATA JOINED IN AUDIOVISUAL REPRESENTATION: BENEFITS AND ISSUES IN ANALYSIS

THORBJÖRN SWENBERG, PH.D.
DALARNA UNIVERSITY
TSW@DU.SE

## ABSTRACT

Participatory methods in design or innovation research can require several different kinds of data in order to cover the object or process of interest to the study. This paper addresses the potential and difficulties of joining and analyzing several data categories into a unified observational video, where all data is represented in either audio, visual or audiovisual form. The core problem of such a complex body of data is that the analysis requires efficient tools and/or methods to match the data complexity. The current case regards a participatory study on film editing, its method and data complexity, and discusses analytical approaches.

## INTRODUCTION

Research methodology is at the core of participatory design (Spinuzzi 2005). Since the study of participatory design or innovation necessarily includes social interaction as well as activities with design or innovation intentions, video (including sound) is appropriate for collecting research data of such events (Knoblauch & Tuma 2011). However, such data is often complex, which entails consequences for the analysis. In order to contextualize actions and interactions further, data can be required that cover work procedures and routines, human-computer interactions, software interface design, as well as levels of detail in the executions of tasks (Hindmarsch and Tutt s.59 2012), not the least regarding aspects of the actual visual thinking that takes place during design and innovation processes (Swenberg 2017, Ware 2004, Wikström 2013). The topic of this paper is how to further develop the methods for analysis of such complex data, to improve the efficiency of the analysis whilst also maintaining the quality of the results. These issues are imperative for researchers' desire to seize the richness of the data and to extract further results from it. The problem is how to extract as much data as possible from participatory videos, including data categories that are yet hard to capture, or to combine in the analysis with analytical means that are convenient and efficient for the researcher. The complex data set I approach in this paper, is captured on video (including sound), and the current aim is to point out some benefits of complex audiovisual research data, as well as the necessary capacities that a new analytical framework for such data must comprise.

Previous studies on participatory design display complex methods, as well as subsequent complexities in the data analyses (e.g. Grenha Teixeira et al. 2017). The participatory case presented and discussed below, from the research project "The Editor's Visual Intention and the Viewer's Visual Perception" also exhibits such data complexity derived from its mix of methods (Swenberg 2016). In this project a film editor participated in a study on whether perceptual precision matters or not in film editing. The editor's participation consisted in making a three minute documentary film sequence, under ecologically valid conditions, in order to examine the role of (audio-) visual perception during the design of film edits. This investigation also included eye tracking of film viewers watching the film sequence made. I will here address analytical issues and requirements for possible solutions that will facilitate the use of rich and complex audiovisual data in participatory research.

## KNOWN ISSUES IN DATA ANALYSIS OF PARTICIPATORY RESEARCH

The first aspect to mention is the interaction designer-user that is both challenging to capture and understand, according to (Grenha Teixeira et al. 2017). Their study explores a large and manifold dataset with a variety of representations, including visual. Particularly, they mention that the iterativety of the design process brought specific issues, since their project was interdisciplinary, and involved several stakeholders. Interdisciplinarity and iterations also contribute to the complexity of the

data of the current case, and thus reveal a similar challenge.

The Living Lab methodology, as discussed by Dell'Era and Landoni (2014), is based on real-life testing and experiments in which users are invited to knowingly participate in a development process. This methodology is accompanied by two principal issues that affect the data analysis in such research designs: (1) the recruitment of participants could be on voluntary basis (open), or by selection (closed), which biases the off-set value of the participants; and (2) the choice of technological platform to use, which affects how creative the outcomes of the design process are. Video recording is a possible, but not necessary, part of the Living Lab research process. The technology aspect of how to comprehend the tool-usage, aligns with the case discussed in this paper, and contributes to its data complexity.

Video recordings display events that unfold sequentially over time, events that include sounds as well as visual information. Therefore, the problem of time relations is an analytical issue, as are the relations between kinds of data that are presented in visual, audio, and audiovisual form respectively (Knoblauch & Tuma 2011). In my participatory study, the specific interest is in the relations between several different kinds of (audio-) visual data (of different epistemology) that occur in various forms. Further, the relations between the verbal and mere sounds made, as well as body language, and actions are aspects that enrich, but complicate the analysis in participatory research (Hindmarsch and Tutt 2012). Such issues occur in my case as well, and the challenge they bring to the analysis is that the data is played out in both time and space, and, when addressed accordingly, they reveal further and more complex results that relate to both audiovisual thinking and thinking in action (cf. Schön 1983). The concepts of *audiovisuality* as three-folded (audio, visual, and combined) in video, *meaningful expressions* as verbal, sound-making, and bodily, *technology* as tool, and interface, as well as participant-researcher *interaction* are re-used in this paper.

## METHODS AND DATA CATEGORIES

During the participatory study discussed here, I observed a film editor along the process of editing a three-minute documentary film sequence. The film editor acted as a participant in my study on design perception in film editing, since she was engaged in the research ends, and the editing task was formulated in mutual understanding between us. These observations took place in the editor's editing suite, where a web camera with microphone recorded me and the editor. The camera image was screened on the editing computer's monitor, covering work as well as social interaction. This monitor was in turn recorded continuously, with all screen output (software interface, embedding the film material being edited, as well as web-camera), and also captured the sounds processed by the computer (including web-camera sound), as well as all key-strokes and mouse-operations (Figure 1).

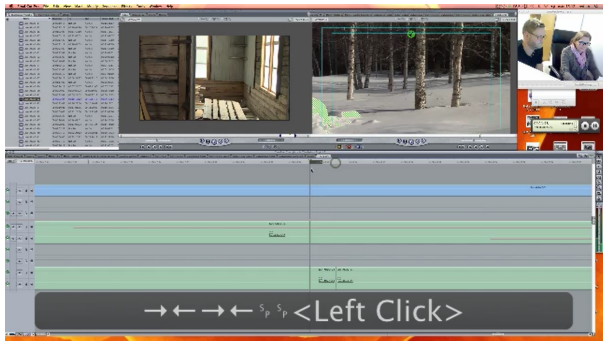


Figure 1: Still frame from a screen-recording video that captures: film material and emerging film sequence, film editing software interface, key-strokes and mouse-operations (white graphic overlays on top of the software interface), research observation (web-camera, upper right screen corner), as well as all associated sounds (see Swenberg, 2016). An excerpt of the video is found at https://vimeo.com/214633391/9a611cfbc1.

These data categories unveil utterances and behaviour from the researcher-editor interaction, what work tasks the editor performs, how she uses the editing software (functions, procedures, and routines), as well as how she treats the audio and video material. After the editing was completed, we had an elicitation session, where the editor expressed her thinking on the particularities of the film material and the edits made.This session was also recorded using the same method. At this point, the engagement of the film editor ended. In total, these continuous screen recordings provide rich data for analysis.

In addition, the completed artefact, a film sequence, is analyzed for perceptual phenomena, and used as stimuli for eye-tracking, along with an altered version of the sequence, to render eye-tracking data. The purpose of eye-tracking viewers watching these sequence versions is to assess the design of the film edits. Eye-tracking data provide measures of to what extent the participating film editor's intentions are fulfilled, as well as whether perceptual precision in the design of the edits matters. Here, it is rather the complexity of the generated data that needs attention, than the particular aim, purpose and method of the presented study. A full methodology for the study is presented in Swenberg (2016), and represented by its instances in Figure 2. Methods for data analysis are addressed below.

## CURRENT ISSUES IN DATA ANALYSIS

The complexity of the data in this film-editing study provides opportunities for many analytical approaches.

The first issue to address is how to separate the different audiovisual data categories of the complex data set, considering their respective audiovisual representation, and forms of measure in analysis.

The audio aspects include (1) the making of a soundtrack to a documentary film sequence, (2) recordings of editor-researcher conversations during the observations and our viewing of the film sequence, (3) an analysis of the recordings of editing (design) work and conversations, as well as (4) audio perception aspects (Figure 3).

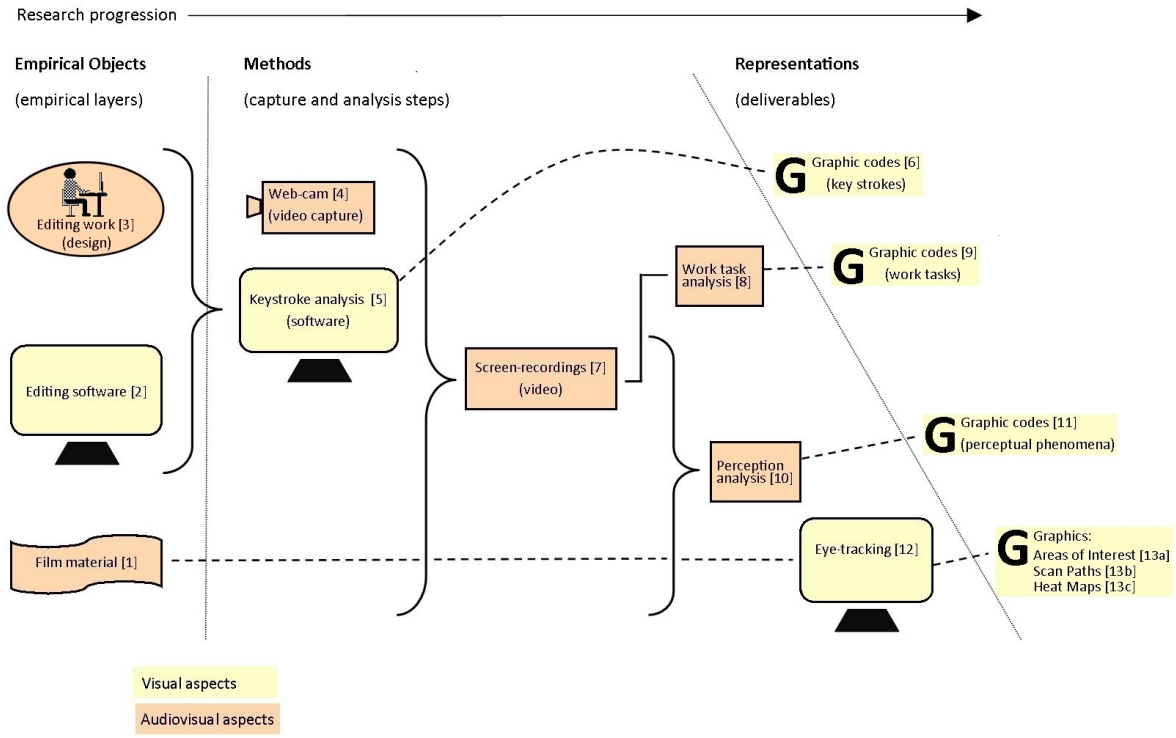


Figure 2: Visual and audiovisual aspects in progression of a research project studying film editing as audiovisual design. Numbers [1-13] indicate the order of occurrence of the different aspects of the project. Documentary film material [1] is handled by means of the editing software [2], which is the film editor's tool. Tool usage [3] and behavior is captured [4+5], and recorded together with the events on the editing computer screen [7], where the film material emerges. Work events are analyzed [8] as well as considerations of perceptual phenomena [10], which are also regarded in direct relation to the completed film sequence. Versions of the film sequence are thereafter screened for viewers, who are eye-tracked [12]. Each analysis delivers some kind of graphical representation of results [6, 9, 11, 13] (from Swenberg, 2016).

The visual aspects include (I) the editing of the documentary film sequence, (II) a continuous screen-recording of the editing software, (III) computer keystrokes, and (IV) recordings of the editor at work (with me as observing researcher next to her) containing video images of the two. The analysis was conducted using this video file onto which codes were added in two steps: (V) a work-task analysis, and (VI) an analysis of perceptual phenomena, using video graphics.

After this analysis, (VII) the two versions (the editor's and the altered one) were used as stimuli in a session where (VIII) viewers were eye tracked. Eye-tracking data was (IX) analyzed according to gaze hits in Areas-of-Interest (AoIs), saccade frequency, and pupil dilation, all occurring shortly after the edit points. Finally, the eye tracking data was presented graphically as over-lay upon the respective versions of the documentary film sequence. Depending on analytical purpose, there are different forms of graphic representation to choose from.

Altogether, in this study, analyses were largely conducted manually, which was time-consuming (with exception for some of the eye-tracking data which was analyzed by the use of software). Data instances, and analyses, are appointed as building upon each other in Figure 3. However, the richness of the data does invite to further analyses, where the relations between data of different categories could be exploited. In the current study, results were only rendered by category, and conclusions were therefore limited. The data categories (1) and (2) thus consist of audio data, whereas categories (I)-(IV), and (XIII) are visual data. Audio analyses are numbered (3) and (4), whereas visual analyses occur in several steps: (V)-(VII) and (IX).

The second analytical issue regards epistemology and what kind of knowledge is possible to draw from each data category, as well as from combinations thereof. During the current study, based largely on manual analyses, the data categories and analyses explicated in Figure 3 were organized as empirical objects, how these objects materialize, what audiovisual indicators they are

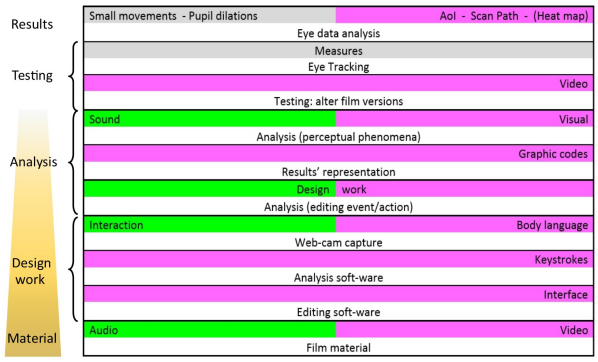


Figure 3: Audiovisual representations of data categories and analyses, building on each other: data categories consist of film material, recorded design work, and eye-tracking testing. Analyses of these data categories, and representations of results, emerge from each data category (developed from Swenberg, 2015). Audio aspects are green, and visual aspects are magenta, whereas other data is grey.

represented by, as well as what kinds of knowledge they support (Table 1). These are (possible) categories of knowledge, addressable by research questions, of which some were explored during the presented study.

However, further sets of research questions would be possible to address with analyses that could cross-read data categories, and thus combine different kinds of knowledge in their respective answers. The point to regard here, is this possible extension of results and further outcomes, concealed in the rich and complex data at hand and hard to extract without the assistance of proper software that can single out, and compare all the available data categories. Examples of such data categories are different kinds of movements, movement trajectories, movement intersections, tonalities, and energies in sounds. These data categories could be combined in analysis with other kinds of already accessible data. However, some of those data categories are only available today through specific kinds of incompatible software. A software package that could, for example, extrapolate eye tracking data over movement trajectories, and compare such data across a time-and-space span, could reveal visual perceptual relations to movement. Such an analysis could give extended results, that add to the understanding of what visual thinking is going on during what actions.

## POSSIBLE SOLUTIONS TO ANALYSIS ISSUES

Nevertheless, software tools for data analysis have become more available for complex sets of data, participatory or otherwise, sometimes derived from mixed methods consisting of both qualitative and quantitative data (Kuckartz 2013). There are several packages of video analysis software on the market that are suitable for observational videos. These software products have tools for coding, annotation, and organization of data sets, as well as for arranging the coded data during analysis, and linking different kinds of data to each other. Parallel qualitative analyses of both audio and video are often possible (e.g. in Transana), and some can also import quantitative data from other sources (e.g. ELAN, and MAXQDA). Analyses can be made presenting many data categories and relations by visual features, and some softwares provide statistics, if desired. Generally, though, any coding of data is made as associated text linked to an audio/video timeline.

However, none of the software I have encountered can address the kind of complexity at stake here, as when visual phenomena need to be addressed visually (not through associated and written codes), in such a manner that visual features in a moving image (e.g. a hand movement) is appointed within the image frame, not next to it. Various visual features within the image

| **Empirical Object:** | **Materialization:** | **Indicator(s):** | **Knowledge supported:** |
|---|---|---|---|
| Film material [1] | Screen-recording video [7] | Software graphics [2] | Film editor's treatment of the film material (AV) |
| | Film sequences [1] | Self-representation | Perceptual (visual and audio) qualities of the material |
| | | Analysis graphics [11,13] | |
| Editing software [2] | Screen-recording video [7] | Software graphics [2] | Software functions used, order of procedure, (usability aspects) |
| | Key strokes [5] | Analysis graphics [6] | |
| Editing work [3] | Screen-recording video [7] | Software graphics [2] | Editing work category, software functions (used and) preferred, treatment of the film material (AV) |
| | | Analysis graphics [9, 11] | Editing work category, perceptual considerations (AV) |
| | | Film material [1] | Material qualities considered and assessed (AV) |
| | Key strokes [5] | Analysis graphics [6] | Software functions used, work precision, order of procedure |
| | Web-cam video [4] | Sounds, movements | Editing thinking (incl. perception) (AV) |
| Viewer gazes [12] | Eye-tracking data | Eye-tracking graphics [13a,b,c] | Gaze behavior |
| Video observation (as Meta level) | Screen-recording video [7] | Self-representation | Film editor - researcher interaction (AV) |

Table 1: Categorization of visual and audiovisual aspects of a research project studying film editing as audiovisual design (expanded on Swenberg, 2016) as represented in Figure 2: empirical objects of study, forms of materialization, indicators, and the respective epistemological outcomes. AV indicates knowledge in audiovisual form. Possible cross-reading of these different categories of data can support richer outcomes if such analyses are enhanced. Thus, further results could be achieved than those addressed in the presented study.

frame stand in relation to each other, and these relations need visually sustained coding that endorses their visual ontology (e.g. trajectories). Still, the usual appointment function must prevail, so that the software's analytical assets resemble the ones used when codes are linked to a timeline. Visual features must be possible to analyze in their visual forms, as well as textually codified.

## DISCUSSION

If video recordings of various kinds are used as a core for participant research, in design, innovation, and otherwise the analysis of video material must stretch beyond the onsets of including body language in the communication processes. Visual thinking and action take place in design events that unfold on-screen in video recordings, both in regular camera-recordings and in videos from screen-recordings. Such events must be possible to analyze on visual terms, as meaningful expressions, using for example graphics upon a video image, where the graphic code is also connected to the analytical functions of a software. These features take audiovisuality as a departure point for data configuration in the analysis, but are missing in today's research analysis software. Another possibility that would emerge with such new functionality, is the possibility to link external, imported data to the analysis, and thus enable the comparison of a variety of data categories, such as tool usage and eye tracking, or user interfaces and iterations.

## CONCLUSION

This paper addresses the possibilities and problems with joining several data categories into a unified observational video, where all data is represented in either audio, visual or audiovisual form. My suggestion is that observational research situations, such as those occurring in participatory innovation or design, can benefit utterly from the richness of the represented data. However, there is a need for efficient tools and/or methods that are capable of providing support for the analyses of rather complex audiovisual data. Visual aspects like movements, trajectories or intersections, as well as audio aspects like tonalities and energies, must be possible to include in non-manual manners, and also included in combinatory analyses. A solution of this issue would facilitate participatory research that uses video observations and screen-recordings, also combined with other data, and lead to more efficient approaches and methods.

## ACKNOWLEDGEMENTS

The moving image editor who participated in this study, Ninni Jonsson Wallin, has credentials in fictional film, documentary film and TV, and holds a position at Dalarna University as senior lecturer in the Art of Moving Image Production, researching documentary film storytelling. The research project "The Editor's Visual Intention and the Viewer's Visual Perception" was supported by the Municipality of Falun, and the European Regional Development Fund.

## REFERENCES

Dell'Era, C. and Landoni, P. 2014. Living Lab: A methodology between user-centered design and participatory design. *Creativity and Innovation Management, 23* (2), pp. 137-154.

Grenha Teixeira, J., Patrício, L., Huang, K-H, Fisk, R. P., Nóbrega, L. and Constantine, L. 2017. The MINDS method: Integrating management and interaction design perspectives for service design. *Journal of Service Research, 20* (3), pp. 240-258.

Hindmarsch, J. and Tutt, D. 2012, 'Video in Analytic Practice', in S. Pink, Advances in Visual Methodology. London: SAGE, pp. 57-73.

Knoblauch, H. and Tuma, R. 2011. Videography: An interpretative approach to video-recorded micro-social interaction. In E. Margolis & L. Pauwels (Eds.), The SAGE handbook of visual research methods. London: SAGE, pp. 414-430.

Kuckartz, U. 2013. Qualitative text analysis using computer assistance. London: SAGE.

Schön, D. 1983. The reflective practitioner: How professionals think in action. New York: Basic Books.

Spinuzzi C. 2005. The methodology of participatory design. *Technical Communication 52,* (2), pp. 163-174.

Swenberg, T. 2015. 'Editorial intentions and viewer perceptions : Audiovisual methods for studying film editing and eye movements', *4th International Visual Methods Conference*, Brighton, UK.

Swenberg, T. 2016. Editorial intentions and viewer perceptions: An audiovisual methodology for studying film editing and eye movements. *Visual Methodologies: A Postdisciplinary Journal, 4*, pp. 13–33.

Swenberg, T. 2017. Framing the Gaze: (Audio-)Visual Design Intentions and Perceptual Considerations in Film Editing. Doctoral thesis. Mälardalen University, Sweden.

Ware C. 2004. Information visualization: Perception for design. San Francisco: Morgan Kaufmann Publishers.

Wikström A. 2013. Storyboarding: framing and reframing opportunities in the front-front end of innovation. Doctoral thesis. Mälardalen University, Sweden.