FRAMING THE GAZE

(AUDIO-) VISUAL DESIGN INTENTIONS AND PERCEPTUAL CONSIDERATIONS IN FILM EDITING

Thorbjörn Swenberg

2017

School of Innovation, Design and Engineering
FRAMING THE GAZE
(AUDIO-) VISUAL DESIGN INTENTIONS AND
PERCEPTUAL CONSIDERATIONS IN FILM EDITING

Thorbjörn Swenberg

Akademisk avhandling

som för avläggande av filosofie doktorsexamen i innovation och design vid
Akademin för innovation, design och teknik kommer att offentligen försvaras
fredagen den 15 september 2017, 10.00 i Raspen, Mälardalens högskola, Eskilstuna.

Fakultetsopponent: Research Fellow Pia Tikka, Aalto University

Akademin för innovation, design och teknik
Abstract

The theme for this doctoral thesis focuses on how a film editor strives actively to perceptually frame and direct the film viewer’s gaze across film edits. The editor’s (audio-)visual intentions concerning the film viewing correspond to perceptual considerations that the editor makes during film editing.

Film editing in this thesis is considered as a kind of design work, foremost motivated by the possibilities of many new approaches, questions, and answers that a design perspective brings, stretching well beyond what previously adopted perspectives have done.

The research questions asked, as well as the presented results, are discussed with regard to design theory, established film production field knowledge, and perception research. Apart from considering audiovisual perception in a film editing context, film editing is also discussed as audiovisual design, and tentative implications for the role of perception in other kinds of design work are pointed out.

The results of my analyses are that there is reason to consider parts of film production as design work; that properties of an audiovisual material affect the creative space; that perceptual considerations are a significant part of the film editor’s treatment of the audiovisual material; that film viewers’ response to film edits varies with the shape of the edits; and that this variation is possible to link with the film editor’s intention regarding the edits, as well as to the degree of fulfillment of the intention.

I conclude that perceptual precision at the edit point is decisive for attaining a desired film-viewing decoding of a film sequence, and that when the perceptual precision is low, it is likely that the perceived continuity of the film fails.

The contribution of this doctoral thesis is that it drives the investigation of actual appliance of perception as audiovisual knowledge in film editing. Hence, it also adds to the more general discussion on perception as part of audiovisual thinking, and how audiovisual knowledge is formed.

The thesis contributes to the area of Innovation and Design through its mix of methods, since the main study considers the creation of new artefacts, the thinking going on during that process, as well as human responses to the artefacts. Conclusively, the thesis provides a thorough example of how a design research perspective can add to the understanding of film production, and its trades and activities.
Acknowledgements

Completing doctoral studies and a PhD thesis is a rather extensive endeavour. It takes effort and support, critical scrutiny and encouragement, in order to go all the way. I have had the privilege of experiencing these crucial components during my process of becoming a researcher, which started out September 2008, and now ends, nine years later.

My sincere thanks go to my supervisors, Professors Yvonne Eriksson and Árni Sverrisson. You believed in my quest, paved the way, and accepted me as a doctoral student. I value your skills as devoted coaches and thorough readers. Dalarna University has sponsored my doctoral studies, for which I am entirely grateful. Pro-Vice-Chancellor Kerstin Öhrn, and Head of School of Humanities and Media, Maria Görts, played a key role in creating the essential conditions for my sponsored part-time leave from teaching. Anyone else who, without my knowledge, contributed in this respect also deserve my gratitude.

Additional sponsoring for my research projects has come from the Municipality of Falun, as well as from the European Regional Development Fund. Participants and companies that took part in my research projects have my deep and humble appreciation for sharing your experiences. My fellow doctoral student and colleague, Per Erik Eriksson, with whom I have had recurring research cooperation, has been a constant inspiration, and devoted much time to chasing data, as well as involving in intriguing discussions on academic matters. Thanks, Pelle, for thought-provoking input! I owe you.

A researcher is not made without an academic milieu where thoughts are shared and challenged. Two higher seminars provided this environment for me: the Audiovisual seminar at Dalarna University, which has become my intellectual home during these doctoral study years, led by Professor Árni Sverrisson, Docent Tomas Axelsson, and PhD Johnny Wingstedt, with its strong attendance of some 35 colleagues, has offered as much critical thinking as support and engagement. I appreciate each one of you! A special thanks to Senior Lecturer, Ninni Jonsson Wallin and Professor Emerita, Sylvia Ingemarsdotter, for contributing input and disagreement. This research group is part the DU research profile of Intercultural Studies, led by Professor Jonas Stier, (formerly Culture, Identity & Design led by Professor Bo G Jansson), which sustained my doctoral studies throughout.

The Information Design Research Group at Mälardalen University, led by Professor Yvonne Eriksson, is likewise an academic milieu of the greatest importance for my development as a researcher. I especially want to mention
Jennie Andersson Schaeffer, Koteswar Chirumalla, Per Erik Eriksson, Ulrika Florin, Peter Johansson, Jonathan Lundin, Inger Orre, Carina Söderlund, Anders Wikström, and Åsa Öberg, as generous colleagues and fellow doctoral students on the path of academic aspiration. I value the atmosphere of possibility, encouragement and constructive criticism you nourish. Very special thanks go to PhD Anna-Lena Carlsson and PhD Anette Strömberg, who led the course in Scientific Writing, and gave me the tools necessary to abandon my confused text constructions, and gain more clarity in my writing. Since then, I started to get published. This group is part of the Mdh research environment Innovation and Product Realisation, IPR, led by Professor Anders Fundin, which has widened my perspectives and made me realize how complex knowledge can be, and thus the research methods required.

I also want to show my appreciation to some people who have been of great help during my research and publishing. First, I want to mention the librarians at DU, for providing fast and accurate service. I thank the many readers of my research papers for their time and constructive feedback that very much improved my texts. Reviewers of my published and presented papers, whoever you are, you have my deep gratitude. My English proofreaders, Professor Irene Gilsenan Nordin, PhD Cand. Joakim Hermansson, and doctoral student Megan Case deserve distinct credit for the readability of my published manuscripts. In addition, Joakim, thanks for the golf rounds and kayak trips that refreshed my spirit, as well as offering intellectual detours into unknown territories! I hope I can make it up to you. Last in this line, by not least, I want to thank Docent Rikard Lindell for a thorough reading of the 90% draft of this doctoral thesis manuscript. I just hope I have done justice to your constructive comments.

Doctoral studies also influence one’s private life, and I have had the wonderful support from my wife and family. Susanna, you are the most flexible and accommodating person I know, and you have helped me get through tough patches. You agreed to my entering doctoral studies, had confidence in me all the way, and I am happy to share this moment with you. Rebecka, Sebastian, och Natanael – ni har hejat på mig och är stolta med mig nu. Tack för att ni finns!

Falun, juni 2017

[Signature]

Thorbjörn Swenberg
Dedication
For Rebecka, Sebastian, Natanael
and Susanna
List of Papers

This thesis is based on the following papers, which are referred to in the text by their ordering capital letters.

A. Swenberg, T. & Eriksson, Y. (Forthcoming) Film and television production as audiovisual design. (Accepted by The Design Collection).


Reprints were made with permission from the respective publishers.
Contents

Introduction ........................................................................................................................................... 3
  The silent knowledge of film editors ................................................................................................. 5
Aim and Objective ............................................................................................................................... 8
Research Questions ............................................................................................................................ 9
Disposition .......................................................................................................................................... 9
Background .......................................................................................................................................... 10
  An outline regarding “perception” within Design Research ......................................................... 11
  Views on Perception ......................................................................................................................... 13
  Perceptual Precision when designing ............................................................................................. 14
The Researching Subject .................................................................................................................... 16
Research Basis ..................................................................................................................................... 18
Delimitations ......................................................................................................................................... 20

Theory and Related Research ............................................................................................................ 21
  Related Research and Core Concepts .............................................................................................. 21
    Design Research ............................................................................................................................... 21
    Academic Media Production Literature ......................................................................................... 27
    Perception Research ......................................................................................................................... 31
    Applied Sets of Concepts ............................................................................................................... 38
    Model Application and Interplay ...................................................................................................... 54

Methods .................................................................................................................................................. 59
  Research Design ............................................................................................................................... 61
    Three Studies – Methods Outline and Interrelation ....................................................................... 62
Data Categories ..................................................................................................................................... 69
  Propositions ....................................................................................................................................... 69
    Audiovisual material ......................................................................................................................... 70
    Actions ............................................................................................................................................ 72
    Eye data .......................................................................................................................................... 72
  Operationalization of Theoretical Concepts ..................................................................................... 74
Methodology and Reflections .............................................................................................................. 77
  The Literature Study on Film and TV Industry Crafts ..................................................................... 78
  The Design Process Management Study ......................................................................................... 79
  The Study of a Film Editor’s Intentions and Viewers’ Perception .................................................. 80
Results .......................................................................................................................... 87
Principal results .......................................................................................................... 87
Result 1: ...................................................................................................................... 88
Legitimacy for studying film and TV production as design ...................................... 88
Result 2: ...................................................................................................................... 89
The material’s effect on creative space ...................................................................... 89
Result 3: ...................................................................................................................... 90
Use of perception in film editing ................................................................................. 90
Result 4: ...................................................................................................................... 91
Viewers’ perceptual reactions to film edits ................................................................... 91
Result 5: ...................................................................................................................... 91
Editing intentions in film editing ................................................................................. 91
A short cross-reading of the results .............................................................................. 92
Detailed results ............................................................................................................. 93
Film and TV production addressed as design ............................................................ 93
Digital materials and constraints on creative space ................................................... 95
The film editor’s use of perception when editing film ................................................ 96
Viewers’ gaze responses to perceptual precision in film edits .................................. 97
The film editor’s intentions when editing film ............................................................. 100
Addressing the results ................................................................................................ 101
Analysis ....................................................................................................................... 103
Analysis addressing the first Research Question ..................................................... 103
Analysis approach and results considered ................................................................. 104
Analysis – step 1: film editing as design ..................................................................... 105
Analysis – step 2: handling of audiovisual material ................................................... 106
Analysis – step 3: care for perceptual aspects ............................................................. 107
Analysis – step 4: viewers’ perceptual decoding ........................................................ 108
Analysis addressing the second Research Question ................................................ 110
Analysis approach and results considered ................................................................. 110
Analysis – step 5: editorial audiovisual intentions ..................................................... 110
Discussion and Conclusion ......................................................................................... 113
In answer to the research questions ........................................................................... 113
Some implications of these answers .......................................................................... 116
Aesthetic problem-solving in film editing ................................................................. 117
Multimodality and audiovisual transients ................................................................... 119
Materiality concerns and consequences .................................................................... 120
Bridging perceptions ................................................................................................... 121
Film editing and perception in design ........................................................................ 123
Contributions ............................................................................................................... 124
Conclusion .................................................................................................................... 125
References .................................................................................................................... 127
Appendix 1 ........................................................................................................141
  Tables ........................................................................................................141

Appendix 2 ........................................................................................................143
  Papers ........................................................................................................143
Film concepts

*Film* is defined as the medium consisting of sound and the stream of still images that move by means of a fast-subsequent exposure of frames (usually, 24 or 25 frames per second). (Paper F)

*A film* is demarcated as a complete object of communication, consisting of film. (Paper F)

*Video* is the material on which film is recorded. (Paper F)

*An image* refers to a moving image with overt features, and objective properties, such as color coherency, aspect ratio and frame rate. (Paper F)

*A shot* is a particular instance of a moving image defined by its own composition, framing, and a start and an end, in accordance with how the camera records its frames. The shot starts where the camera starts recording, continues through that recording, and ends where the recording stops on that specific occasion (Papers E and F).

*An edit* refers to every visual shift in a sequence, where one shot ends, and is followed by the start of the next shot. (Paper F)

*Edit point* is the point in time (from the film’s start, or on a film editing tool’s timeline) where an edit occurs. (Paper F)

*Incoming shot* is the shot running until the edit point. (Paper G)

*Outgoing shot* is the shot starting at the edit point. (Paper G)

*A transition* is the audiovisual phenomenon occurring at an edit, as a shift between two shots. Transitions may appear as a smooth blending of shots, in some other graphical form, or it can be immediate. (Paper F)

*A cut* is an immediate transition, from one frame to the next, between two shots. (Paper F)

*Jump cuts* are cuts between shots where either the focus object changes, while the background stays constant from one shot to another, or vice versa. (Paper F)

*A sequence* refers to when several shots are joined together into an order, where there are no gaps between them, and where a sound track runs along those shots. (Paper F)
*Editing* is the work that a film editor does, including choosing shots, ordering them into a sequence, removing, adding, ordering sound, and trimming the transitions between shots. (Paper G)

*Film perception* is used to encapsulate human capacities to adapt natural perceptual functioning to the experiencing (see-hearing) of a film (Berliner & Cohen 2011, Smith 2012, in Paper G).

*Continuity* refers to the human experience of the world as stable, i.e. continuous in its existence in regard to our sensory input (Hecht & Kalkofen 2009, Smith 2012, in Paper G).

*Perceptual continuity* is the viewer’s experience of a film as presenting a continuous flow of un-broken audio-visuality. (Paper G)

*Film continuity* is the audiovisual construction of filmic expressions meant for perceptual continuity. (Paper G)

*Discontinuity* refers to any sudden change in the world that causes sensory transients that are detected (Smith 2012, in Paper G).

*Perceptual discontinuity* is defined as viewer perception or attention to audiovisual transients that disturbs the experience of a continuous audiovisual flow. (Paper G)

*Film discontinuity* is constituted by the filmic construction provoking the above transients. (Paper G)

*Continuity edits* are transitions between shots, made to appear smooth or invisible to the eye, where a film sequence supposedly appears as one continuous shot, without breaks. (Paper E)

*Discontinuity edits* are cuts between shots made with transitions overtly transient, for the viewer to respond to perceptually, attentionally or consciously. (Paper E)
Introduction

There were two shots of the same everyday event, a bartender pouring a pint of lager from a tap, shot from different angles, which I joined in sequence while editing a travel show for television in the 1990s. Immediately after joining the shots, I could not see the cut between them – neither could I see them while running the sequence forwards, or backwards, or even while running it in slow-motion. The rapid and curved movement of the bartender’s swinging of the pint from the tap, up to the counter, completely fused the cut between the two shots. Ever since, I have been fascinated and puzzled by our visual perception of moving image edits. How come some edits can be so disguised before our eyes that not even a trained eye can perceive them?

Moving images are commonly used to convey a story, but may also be exploited expressively. How moving images are perceived by viewers affects how they are read and related to context. Any production processing of a moving image can influence how it is perceived. Perception, and how it concerns the design of film edits, is the topic of this thesis. Film editing is engaged in directing the eye to key features of moving images, as a form of visual communication medium, by narrowing down alternatives for attention. Since film is a multimodal medium, film perception also includes the interaction between audio and visual features. These matters are here considered as design issues requiring more thorough understanding, addressed from a design research perspective.

In my attempt to approach how perception is engaged in film editing, and whether perceptual considerations matter to film viewing, I use two objects of study: a case of observing documentary film editing (vimeo.com/214633391/9a611cfbc1), and an eye tracking of viewer responses to edits made during the editing. Documentary film is chosen because it is more challenging to edit than fiction film, in the sense that it is harder to achieve an audiovisual flow with documentary footage. In addition, the role of audiovisual perception may influence editing intentions about film decoding, considered as balancing of both top-down, and bottom-up attention. Perhaps these issues are relevant, and transferrable, to other design contexts where (audio-) visual perception is of interest.

Design science concerns “the study of the principles, practices, and procedures of design” (Cross 2001), and employs “a federation of sub-disciplines” (Gasparski & Strzalecki 1990) in understanding various facets of the field of
design. Perception psychology and cognitive science can be consulted as supportive disciplines for explaining issues of attention in design. Design psychology and design cognition exist as fields of design research (Horváth 2004). Nevertheless, studies that specifically address whether the gaze gets framed, or not, by visual design in various contexts amount to only a few dozen.

This thesis raises issues that are not usual design or innovation research concerns. Topics where “visual design intentions” are paired with “perception” within design or innovation research occur only in a hand-full of texts. Audiovisual design has only seldom been attended to within the contexts of design, animation, human-computer interfaces, or media production. Likewise, there are just a few research results on the role of perception during design work presented in journals on design, architecture, or on the sociology of work and occupations. The current approach is underrepresented.

The Background chapter is therefore devoted to a brief sketch of the place of this thesis within design research, and its relevance to that field of study. In addition, in order to indicate a perspective, my understanding of the concept of perception is given some preliminary space in that chapter. The disposition and structure of the thesis will be presented and explained under its own subheading.

My Contribution with this thesis specifically calls for a deeper understanding of film editing, its relation to human perception, and how perceptual precision may be employed by a film editor in the form of aesthetic problem-solving. This understanding contributes to a more overt understanding of the characteristics of the film editor’s trade. I also present new mixed methods used across research fields, as well as additions to the appreciation of audiovisual knowledge and its appliance in communication. In particular, I want to make a contribution to the prospects of considering film and TV production as audiovisual design.
The silent knowledge of film editors

Taken, for now, that audiovisual design can include film and TV production activities, the link between film editing, design and perception needs some preliminary justification.

Previous research has identified visual features that a designer has to attend to that provoke perceptual responses in humans (Ware 2004). Yet, how visual perception is addressed and used by designers can be further understood if given more research consideration. We know only little about how visual perception is employed by a designer in the design situation, and what material processing is specifically subject to perceptual considerations. Just as little do we know about how a more thorough understanding of visual perceptual phenomena could be applied in design work, or how such appliance could affect design processes. In the present study, I will address these problems in the case of documentary film editing. I will do so to exemplify why film editing can be understood as audiovisual design, simultaneously as I address matters of perception in film editing.

Audiovisual design is here ascribed to the activities involved in the creation, altering, and adaptation of sound and image constituents, as well as how to make them interplay with each other in a desired way: e.g. sounds, shots, camera-movements, lighting, and objects within the moving image frame, and their combination in an object of communication (Swenberg 2012). When I use “(audio-)visual” in the text, it refers both to instances that are audiovisual, as well as entirely visual instances. As in other design, audiovisual design addresses the user of an artifact, e.g. a film. But in order to distinguish any user, as a possible or actual individual using any designed object, from possible or actual film-watchers, the latter are from now on referred to as viewers.

A viewer’s experience with moving image editing can be that sometimes edits jerk, and sometimes they pass fluidly before the eye. An attentional theory of continuity editing is a theoretical basis for fluid edits, presented by Tim J. Smith (2005). I find it evasive as to what extent such a theory, or its constituents, is actually applied by moving image editors. The purpose of the current study is to explore whether a film editor does, in fact, regard perceptual phenomena, how visual perception is employed, and what processing is undertaken that transforms perceptual considerations into concrete material form. If perception is considered, as Smith (2005) suggests, it has to involve some sort of precision regarding how the perceptual phenomena appear and evolve during the audiovisual flow of a film. It is also puzzling whether any perceptual considerations made by an editor contribute in any significant way to the completed film. Thus, part of the problem is that the role of perception in film editing needs further clarification.

In parallel, the material designed, film, consists of digital audio and video that materialize as sound and moving images. This material needs to be processed. Editing an existing stock of documentary film material involves
choosing, adapting and altering digital video and audio files by means of editing software. At the end of that process, the artifact is expected to express something that is meaningful and credible to others. The processing of the material will include figuring out the qualities of the material, as well as creating the expression. Digital materials lend themselves open to many trials, iterations, and reconsiderations, but also convey specific limitations (Manovich 2001 & 2012). How well attuned the designing of the artifact is to (audio-)visual perception needs deeper understanding, as a further part of the research problem, in order to contribute to improving the handling of the material.

Understanding perceptual handling of the material can thus bring better ways for designers to solve problems faster and more satisfactorily. Accordingly, designers can spend more time on other creative matters than on concerns regarding how material properties are perceived perceptually. A deeper knowledge and understanding of the role of perception when audiovisual material is processed will thus allow the designer more creative space (cf. Sveberg 2012). Additionally, if audiovisual designers work in a system or network with other designers, the resulting time gain promises positive effects for the overall design process (cf. Johnson 2005, Sveberg 2012, Wang et al. 2007). Thus, part of the research problem regards how perceptual assessment of an audiovisual material matters to the enabling of an audiovisual designer’s creative space.

Further, the adapting and altering activities of film editing must somehow include judgements, outspoken, or reflected upon “in-action” (cf. Schön 1983) that address what audiovisual features are to be kept or discarded in order to make edits appear as desired (cf. Alexander 1964 p.53-55). These editing activities are possibly un-articulated, and thereby un-recognized, audiovisual thinking that preserves film editing in the mystery “cocoon of silence” in which it is claimed to reside (Murch 2006 p.x). If non-verbalized knowledge is disregarded as “mere intuition” or “gut feeling” that lacks reason, film editors’ audiovisual thinking will remain un-empowered, which is a major part of the problem.

If, on the other hand, audiovisual thinking in film editing can be addressed in a similar fashion as the kind of thinking that Colin Ware (2008) points out as careful attentional and perceptual activities in visualization design, it might be possible to more deeply understand the kinds of knowledge employed by film editors. *Visual thinking,* as a concept, is established to denominate the process of visual perception—attention—and evaluation of visual information to use, and possibly combine with other information, or neglect, in order to make sense of visual information. Then, the film editor’s visual thinking might also relate to intentions regarding viewers’ viewing. In the instance of audiovisual design, the uttermost reference for how perceptual phenomena matter is the visual responses to the artefact by viewers. In this thesis, these responses from viewers are related to perceptual precision, as a concept for accuracy in perceptual considerations when an artefact is being designed. Yet
a part of the research problem, is whether perceptual precision is important for viewers’ perception, which requires examination, as well as the relation of perceptual precision to editorial intentions.

A film editor’s intentions are context dependent. This thesis is based on objects of study derived from documentary film editing. The aims that frame an artefact affect the design process. Continuity film editing strives to support the fluid comprehension of a story. Engagement with, and understanding, a narrative requires attention to the details that construct the story. Perception and attention are expected to make an impact on fluid comprehension. How such impact manifests itself, as perceptual reactions of film viewers, is the conclusive part of the problem, addressed here, that needs more investigation.

In sum, I state the problem addressed by this thesis as a lack of clarity regarding how perception is used in audiovisual design, particularly in film editing. This problem expands to include our insufficient understanding of how materials are treated from an (audio-)visual perceptual stance, where the impact of perceptual understanding of materials on a designer’s creative space is not yet illuminated satisfactorily, and needs to be done. Still, the consequences of this undeveloped understanding of (audio-)visual perception in design is the major problem: there is a bulk of un-articulated audiovisual knowledge and skills, related to perception, that designers embody, which could empower them if articulated, here, specifically regarding film editors. This empowerment does not happen unless the role of perception in design is satisfactory researched. And the problem continues: we do not know how perceptual skills, like precision, matters in regard to the fulfillment of design intentions, as in this thesis: film editing intentions; a lack of knowledge that generates a gap in how perceptual and attentional impact on comprehension are manifested, particularly in film viewing. Accordingly, we do not know whether to prioritize, or not, perceptual knowledge, skills and understanding in curricula as well as in practice.

The relevance of the results of this thesis is primarily how we should understand the importance of perceptual precision in the audiovisual design of continuity editing. Secondly, the results contribute to understanding how (audio-)visual perception is utilized by designers, as well as how it is applied in the processing of a digital material. Thereby, some possible effects on, and recommendations for, design processes are highlighted. When such understanding enters design curricula, design thinking and design skills can be more productively developed (Pratt 2011, Costa et al. 2014). Inexperienced audiovisual designers can also improve their skills by assimilating these findings into their body of knowledge.

In society, at large, the increased usage of visual expressions – images in particular – calls for a deeper understanding of visual practice, and for this understanding to be adopted by image professionals as well as the users of vernacular images (Aspers, Fuehrer, & Sverrisson 2004). Within the extended sphere of education, understanding of audio-visuality and audiovisual design
is increasingly important as a sub-set to the area of media and information literacy (MIL) (UNESCO 2014).

This thesis addresses how explicit knowledge regarding the role of perception in audiovisual design work can contribute to an improved understanding of certain perceptual knowledge and skills that audiovisual designers employ, but that can also contribute to our understanding of audiovisual knowledge, as we apply it in communication, on a general level.

**Aim and Objective**

The objective of this thesis is to increase understanding of the role of perception in audiovisual design, particularly in film editing. Audiovisual aspects of design occur frequently, not only in audiovisual media, but also explicitly in digital devices of various kinds, and in software interfaces, as well as in designed physical spaces (cf. Bartholdy 2008). Addressing audio-visualuality in design from a perceptual point of view will add appreciation for the perceptual effects of audiovisual details.

My aim in this thesis is to trace down the use of perception in audiovisual design, and to identify how a documentary film editor, as an audiovisual designer, employs perceptual precision as a means of realizing design intentions of fluidity and continuity.

I understand *design intentions* as what a designer, or design team, appoints as the purpose(-s) of a particular design, design activity, or design decision, as well as how a designed item is expected to be responded to by its user (or viewer). Part of the aim is to link design intentions with efforts to frame viewers’ gazes to what the designer estimates to be crucial details in the audio-visuals designed, with examples derived from documentary film editing. If the designer actually achieves such framing of viewer’s gazes, as a second aim, it may be possible to establish that this framing of gazes also has perceptual effects for the viewers that are important for their understanding.

Design researchers can expand on such results when addressing perceptual issues in other design areas. From a general academic point of view, the thesis also contributes to the understanding of perception as an instance of audiovisual thinking and audiovisual knowledge. In addition, the knowledge derived from this thesis targets designers inexperienced in (audio-)visual design, as well as educationalists in programs where (audio-)visual design plays an essential part (Pratt 2011, Costa et al. 2014). Interweaving perceptual skills with other (audio-)visual design thinking, in curricula, as well as in practice, will reinforce learning throughout the audiovisual design field.

Furthermore, understanding the role of perception in (audio-)visual design is also relevant for established designers who want to learn more regarding (audio-)visual perception. Deeper explicit knowledge regarding the function of perception in design will contribute to the development of their practice.
Research Questions

In order to explore the links between (audio-)visual design intentions and (audio-) visual perception employed by film editors in designing film edits, and to understand to what extent perceptual precision impacts design results, as well as viewers’ and users’ perception and understanding of the artefacts, I use the following research questions in this thesis (RQ):

What does perceptual precision attain in the design of continuity edits and for viewers’ film decoding?

In what ways could (audio-)visual intentions be dependent on perceptual precision in the design of continuity edits?

The second question is regarded as the core question of the thesis. However, it needs contextualization, which the first question addresses.

Disposition

The research questions of this thesis address perception of continuity film edits, and how they are designed audio-Visually. These questions need some further contextualization as the core of the study.

After this brief introduction to the topic of the thesis, and presentation of its problem, aim, objective and research questions, I will next sketch the background of the thesis within design research (including an initial overview of “perception” and a few other key concepts), in relation to this. Thereafter, I will present myself, The Researching Subject, as a doctoral candidate, together with my professional background, followed by a presentation of the empirical studies providing the evidences for the thesis, in the Research Basis section, and Delimitations.

Related research provides much of the theory applied in this thesis, therefore these matters are taken up in a common chapter, Theory and Related Research. Core Concepts are drawn from the areas of design research, media production and perception research, and are thereafter merged into sets of concepts, and further developed as Models Applied in Interplay.

The methods I have used during my research are presented starting with the Research Design of the three founding studies, followed by a discussion of the Data Categories generated, as well as reflections on the methodology.

The emerging Results are then highlighted: first, the principal results in brief, and then the detailed results at length. The Analysis chapter relates the research questions to the theoretical models applied, as well as to the data categories, as a staging of my analytical approach.
Finally, I draw conclusions from the results as answers to the research questions, and engage in a further discussion on their consequences and importance in regard to the presented problem, as well as how this thesis could become a point of departure for others in this field of research and education.

Background

Design and (audio-)visual aspects of design include many instances and several fields of research. How audio-visuality is understood as a part of the thinking that goes on during design is of core interest here, as an outset for this study.

The point of departure for this thesis is Colin Ware’s (2008) claim that perception is an essential part of visual thinking, consisting of a “series of acts of attention, driving eye movements and tuning our pattern-finding circuits” (Ware 2008 p.3). My interest is in the role of audiovisual perception as part of audiovisual thinking, and how it is applied by designers. Studying these matters can provide new knowledge and insights in developing design research, as well as in the fields of design.

There seem to be occasions and situations within film production (Smith 2005) where (audio-)visual perception has apparent links with audiovisual design. In this context, there is evidence of designers having to actively process (audio-)visual perceptual phenomena in order to narrow down and focalize, or frame, viewers’ perception towards such information that is more important, taking into account the purpose of what is being designed (Myers 1989, Suwa & Tversky 2002, Maiocchi 2015). From this point of view, designing means directing the gaze of both the user and viewer (Santella 2005, Ware 2008), which requires some kind of perceptual precision-work, or finesse, in order to succeed.

Visual perception is often mentioned in the design literature (e.g. Lawson 1990, Simon 1996, Norman 1998, Cross 2006, Koskinen et al. 2011, Aspelund 2014, Smith & Schank Smith 2015), yet, it is acknowledged in several different ways. It is either considered as sensory functioning, as in the tradition of perception psychology, or, more commonly, as the act of including sensory input into the mental processes of understanding the surrounding world, as prevalent in the broad scope of academic studies of cultural phenomena:

As the act of organizing sensory information in a meaningful fashion, design can thus be defined as the articulation of perception. Design is always an intervention in the world, and the accumulation of our designed actions has utterly altered the physical reality of our material environment and, consequently, our perception of it. [...] design both articulates and changes perceptions of the world.

(Jen Rhee, “Perception” in Design Dictionary)
This is a bi-directional understanding of perception in the act of designing. We first use perception to organize matter into new configurations as artifacts. These artifacts become an addition to, or altering of, the existing world. Then, we use perception as sensory input to re-interpret the world with its new additions or alterings. This view also assumes our previous perception of the world, which is necessary, in order to apprehend whatever needs there are for additions to, or altering of, that world, before the design of new artifacts.

An outline regarding “perception” within Design Research

Perception is obviously needed when creating new artifacts. The focus of this thesis regards the designer’s use of (audio-)visual perception in steering the attention of user and viewer. The scope of the thesis will primarily focus on audiovisual design, explicitly considered in the form of documentary film editing. First, however, the role and understanding of perception within the design field will be sketched, so that perception can be related to design intentions and design processes. Especially addressed will be the appreciation of visuality within some key works of design research.

When Herbert Simon (1996) discusses *The sciences of the artificial*, his reasoning has its point of departure in the extensive discussion amongst scholars and designers in the 1960s about a scientific basis for, and understanding of, design. Simon is concerned with understanding human intelligence directed at the evolving social world, full of artificiality. Design thinking is at the center of his discussion. Understanding complexities requires (audio-)visual thinking, including perceptual processing of visual representations, to attain solutions to problems. Attention is a human bottleneck resource, and designers should strive for the avoidance of “[...] irrelevant distractions of [users’] attention” (Simon 1996 p.143-144). “Focus of attention is the key to success – focusing on the particular features of a situation that are relevant to the problem [...]” (Simon 1996 p.109). Further research is required, he claims, in order to develop a theory regarding attention and design.

In Donald Norman’s (1998) understanding of design, design starts with perception, to perceive the world in order to understand it, to be better able to affect it. Designers must assume that the human attention capacity is limited, and therefore selective, which urges them to guard against misleading focus of attention when designing. Norman’s concern is outlined in his book *The design of everyday things* (1998), where he regards visibility as one of the most important principles in design. What a user needs to understand about a designed object should be visibly comprehensible, and its visible appearance, “the system image”, should be easy to interpret (Norman 1998 p.189-191). The designer must therefore strive to make all that is important about an artifact visibly apparent, and when this is not possible, he regards audio to be an optional complement.
The most important aspect of design work, according to Karl Aspelund, is to be able to illustrate and explain ideas (2014). In *The design process*, Aspelund claims that sketches can explain things that words cannot, and that they do this quickly as well. He sees “gestalt perception” as a means for the designer to organize visual information, and a function that the viewer uses to cue the information for interpretation.

IDEO’s CEO, Tim Brown, in *Change by design: How design thinking transforms organizations and inspires innovation* (2009), holds visual thinking not only to be the absolute core to design, but also to be crucial for decision making. He maintains that visuals express ideas that cannot be captured by words alone.

Another interesting aspect of the use of visuality and design thinking is proposed in Anders Wikström’s thesis, *Storyboarding: Framing and reframing opportunities in the front-front end of innovation* (2013). Wikström suggests that the combination of design thinking, visual thinking, and narratives through visual representations enhances innovation processes. Visuals expand the reasoning about, and explanation, of ideas (Wikström 2013).

In *Developing your design process: Six concepts for studio*, Albert C. Smith and Kendra Schank Smith (2015) point out to design students that directing perception to the core of a design is a key to avoiding confusion and enhancing understanding. They also stress visual means for communicating design concepts as the only possible way of explaining the concepts accurately: “Emphasizing the parts of your project that are more important helps to ensure that everyone understands what you are trying to accomplish” (Smith & Schank Smith 2015 p.174).

Kees Dorst and Bryan Lawson, in *Design expertise*, point out that visual representation is a vital part of designers’ professional practices, and is used in the process of developing visual thinking on a particular design (2009). “They think directly by manipulating graphical information [...]”, Lawson and Dorst claim (2009 p.104), and refer to the realm of perception psychology to explain how this is possible.

Attention and perception as studied within psychological studies are also recognized by Nigel Cross, in *Designerly ways of knowing* (2006), as aspects of design behavior requiring further study. He considers aesthetic perception as an unmistakable part of design thinking, with reference to neuropsychology (Cross 2006 pp.21). Visual expressions and representations are other elements that design thinking depends on, for instance, for the continuous development of design solutions to problems. Cross claims that the representational medium is critical in enhancing design thinking.

In *How designers think*, it is clear that Lawson (1990) considers designers’ perception to be a core issue. He claims, it is important how things are perceived and attended to, in order to come up with designerly good solutions to problems faced or given tasks. Lawson also includes visual representations in his reasoning in a manner that seems to take them for granted within design
practice. What he specifically addresses is the fact that visual representations often fail to give justice to all aspects that a designer has to consider (Lawson 1990 pp.170), thus adding erroneous detail perception.

Last in this design research survey, I would like to mention Donald Schön’s concept of framing a problem, as a significant starting point for a practicing professional. In his work, The reflective practitioner: How professionals think in action (1983), Schön develops the idea of testing the framing of a problem. This idea will be adopted in this thesis, when the problem of framing gazes for an intended perception is addressed.

Views on Perception

Unmistakably, perception and its understanding are of importance in design research. However, perception is not an undebated concept. In attempts to understand its role in the processing of sensory input into useful information, it has been addressed by scholars from such various disciplines as, for instance, cultural studies, cognitive science, and computer science. Within the field of artificial intelligence, perception has been considered from a sensory-system perspective, as input coupled to action (Cassimatis, Trafton, Bugajska. & Schultz 2004, Froese & Ziemke 2009, Lewis et al. 2011, Sun & Hélie 2013). When human cognition is studied, perception is considered as sensory-input modules related to local environments, allowing humans’ automatic functioning, as brain processes separate from ideas one might have (Pylyshyn 2003, Goldstone, de Leeuw & Landy 2015).

From the perspective of humans as cultural beings trying to understand sensory inputs, perception is considered an information selection process, driven by conscious or unconscious needs to integrate new information (Hamlyn 1996, Crary 1999). “With an admirable economy we learn to see only so much as is needed for our purposes; but this is in fact very little […]” (Fry 1981 p.18). It is not always clear if this understanding of perception differs from the concept of reception, which emphasizes the integration of new information into one’s understanding. From a design perspective, Gibson’s theory of “direct perception”, including objects’ “affordances”, is of particular interest (Gibson 2015[1979], Turvey, Shaw, Reed, & Mace 1981), since designers usually want the items they design to have properties that are easy for the intended target group to grasp and relate to (Norman 1998).

Visual perception has attracted widespread interest within studies of art and visual culture (Arnhem 1954, Mitchell 2005, Pinna 2008), and the integration of perception and cognition is a prevalent view in this field, where sensory information and its processing are regarded as a whole bi-directional chunk of activities. In a cultural context, the concepts “gaze” and “glance” have been used to label different ways of looking at things, with different attitudes (Bryson 1983). Whereas, in the perception psychology tradition, “gaze” is simply the active look of (a pair of) eyes, having a direction (Wade &
Tatler 2005, Land & Tatler 2009), regardless of the purpose, or attitude, of the looking agent. This approach allows for the separation of eye movements from tasks given to the visual system by cognition, and thus links between tasks and eye behavior may be explored.

I make a point of regarding perception and cognition as different addressable processes, even if they are linked and related, and thus not altogether separate. Nevertheless, some perceptual activities are found to exist detached from cognition (Pylyshyn 2003). For a clear reasoning, it is useful to adopt different concepts for different aspects of perceptual and cognitive activities. Therefore, in this thesis:

*Perception* is the process by which sensory organs respond to, or pick out salient information from the surrounding world, and pass this information to the brain for further processing. *Audiovisual perception* is thus a narrowing down of perception with a focus on the visual and hearing senses, and their co-delivery of information for cognition. *Attention* is the cognitive state when “some sensory inputs are processed faster or deeper than others, and thus become more readily available for action, memory, or thought” (Lamme 2003). *Transients* are salient features of the world that stand out and call for attention, e.g. an obtruding sound or image property of a film, which can induce attentional or perceptual reaction in a viewer (Smith 2005). If this takes place, it works as an *attentional cue*. *Perceptual precision* is my construct for the degree to which an ideal consideration of transients is achieved by a designer, either avoided, or exhausted, due to what is perceptually convenient (Paper F).

In the end, these activities might be argued to be one over-arching process of situated cognition, serving the same essential purpose, namely, letting a creature with mind and body interact with its environment (Eriksson 2009). This interaction, as a form of embodiment, can occur in a vast variety of ways, but film viewing only is the focus of interest in this thesis. The design point of emphasis in both these cases is how easily and fluidly members of the target groups perceive the object they face, since they are assumed to use the sensory perceptual input for specific cognitive processes. This perspective relates to Gibson’s direct perception. Designers’ adherences to perceptual qualities of the objects they design are expected to affect how the affordances of the objects are perceived by others. How meticulous, or otherwise, designers’ adherence to perceptual qualities is in trying to achieve an accurate (intended) attention is what I intend by the term “perceptual precision” (see Paper F).

**Perceptual Precision when designing**

This thesis sets out to examine how some forms of perceptual precision function in audiovisual design, and whether it is possible to comprehend to what degree designers’ perceptual precision matters in reaching an intended design outcome. Design ideas and intents depart from the objectives framing an artefact. Central to these matters is how perceptual precision is achieved,
to what extent it is accomplished, and whether the intended target group members perceive the artifacts as intended. Their visual perception is therefore an act that is considered as being possible to measure via, for instance, eye tracking, and what is captured is their gazes, understood as within perception psychology (cf. Wade & Tatler 2005, Land & Tatler 2009).

How an object’s affordances are perceived also depends on the object’s material properties (Gibson 2015[1979], Norman 1999). In this thesis, the material considered for the object of study consists of digital sound and moving images, in the case of documentary film editing. Hence, the perceived affordances are conditioned by audiovisual material properties, and these properties are what designers deal with when they design (Norman 1999). Subjected to these circumstances, (audio-)visual designers can be assumed to have perceptual intentions with the (audio-)visual artefacts and materials they process. However, this is not assumed, but considered in the research questions of this thesis.

How artefacts are perceived is known to qualify as a key factor in customer responses (Veryzer 1995, Norman 2002, Crilly, Moultrie, & Clarkson 2004), or whether it is useful to its users (Tractinsky 1997, Thüring & Mahlke 2007, Sánchez-Franco et al. 2013). Even if the artifact under inquiry in this thesis is conditioned by its specific design intentions that differ from those of customer products or user interfaces of software, it is interesting to see whether the viewers’ perception is affected by the design, since there is evidence that humans respond to their audiovisual perception of materials (Fujisaki et al. 2014). If viewers’ perception is affected, the importance of (audio-)visual designers’ perceptual precision is stressed.

Thus, understanding (audio-)visual perceptual precision will be discussed as fundamental for audiovisual knowledge in the design context. Designers’ visual perceptual precision could possibly be addressed differently as a function of visual thinking during design work. Visual perceptual precision is also possible to address in objective terms, since perceptual phenomena can be challenged, and members of target groups can be eye tracked. Further, understanding of perceptual precision as integral to audiovisual thinking will be suggested as influencing future research, teaching, and further education in the area of audiovisual design. Design research could possibly benefit from understanding perceptual precision in any area that employs visuality. Design curricula could potentially be improved by addressing perceptual precision as part of visual design strategy and applied assets for reaching intended target group perception. Active designers could develop their own skills further, by learning more about perceptual precision, in order to possibly enhance their understanding of their professional habits, and thus improve their fluidness in designing.
The Researching Subject

How to understand the professional knowledge and skills in my previous work as a TV editor was what brought me to pursue scientific research. The problems I present here, what the role of perception is during film editing, whether perceptual precision is important, and how perception and attention become part of a film editor’s audiovisual thinking, is a genuine interest of mine. As a PhD candidate, it is fulfilling to present this thesis as the result and answer to the queries that incited me into becoming a researcher, and which first bothered me as an emergent doctoral student.

This issue echoes the situation highlighted by Murch (2006) that film editing, and the knowledge and thinking it includes, has remained unarticulated throughout the century it has existed. Experienced film editors encapsulate this knowledge, as embodied skills, whereas research on how humans can perceive film edits and make sense of film viewing has been conducted in labs, far from actual film editing. With my experience from both fields, I want to pursue the issue further, in trying to formulate what audiovisual knowledge film editors use, by means of a combination of research methods inherited from different disciplines. This approach is in accordance with the one advocated by Sean Redmond and Craig Batty (2015), where moving image perception research relates closely to its field of practice.

My first paper (Paper D) not only presents my first research results, but also presents the very issue that has concerned me since I worked as a TV editor, namely, how come some moving image edits become so fluid that they even disguise themselves before my own eyes, moments after I have made them? I have been preoccupied with this topic ever since, and with the research approach developed in that first pilot study (Paper D): knowing a film editor’s intentions with edits made (vimeo.com/214673656/3418c4295f), and comparing viewings of versions of a film sequence, that differ in how the edits are shaped. My final paper (Paper G) presents the results of the fully developed study, with 33 participants eye tracked while viewing two different versions of a documentary film sequence, made exclusively for this study (vimeo.com/210944884/be8001e786, vimeo.com/21095506/9ed3cb3b6e). As a lecturer in Moving Image Production, I also see benefits for students in realizing the function of perception early on in their endeavors while creating and processing moving images.

In between writing those papers, I came to realize that understanding design is very much about understanding processes. I have also come to understand the common ground between moving image production and design: the similarity in how to frame and approach issues and problems that occur during production processes, which are often aesthetic problems and issues regarding form and meaning. This similarity convinced me of the benefits of framing the study of moving image production as audiovisual design (see Paper A). Therefore, I have also engaged in a project that addressed new film
and TV production processes as a design research project (see Papers B and C, and my Licentiate thesis, Swenberg 2012). Once more, I expect students, or others, to attain a more robust understanding of their own production processes, as well as others’, by means of a design perspective.

As I returned to perceptual matters of editing, I continued to study moving image edits, now in a more extensive project, where I involved a film editor to edit a sequence for me. This time, I spent more time and effort studying the editor at work, and found that it was necessary to analyze her work and processes to understand how she employed perception when designing the edits. The analysis resulted in Paper F. The drawback of my pre-understanding as a former TV editor is that I do not see all that happens afresh, as new phenomena to analyze. The benefits gained from my previous work experience, on the other hand, are that I can more unconstrainedly follow another’s editing work, understand what kind of work is going on, if there is a constrain of some kind, what tool usage is taking place, and from that address the editor with accurate questions. Thus, moving image editing is not all new to me, and I can more thoroughly examine the work and its process without getting bogged down with unimportant issues.

I continued the study by manipulating the edits from the editor’s sequence, as I made a second version of it, to grasp whether her perceptual precision at edit points mattered. The two versions were screened for viewers who were eye-tracked, and the results are presented in Paper G. Altogether, this project compiled so much data generated by means of several methods that the methodology had to be discussed separately (Paper E). This was done, first, presented at a conference, and then further developed and published in a journal.

I continue to be interested in audiovisual perception as an aspect of design, as well as in combining data categories, and anticipate further research efforts using such approaches.
Research Basis

In order to arrive at how audiovisual perception is employed in design processes I have employed a selection of research procedures. This doctoral thesis is based on three studies and a pilot study that are here presented as the thesis research basis. In the following account, my contribution to each study is clarified.

The three studies and the pilot-study are presented in the following, according to their scope: starting broadly by studying film and TV production as audiovisual design, narrowing down the focus to audiovisual design processes, and the use and influence of perception in design work. Overarching, it is possible to read my research as following two trails: one design process trail that addresses film and TV production as design, in order to investigate the creation and processing of moving images from a perspective of organization, workflows, constraints, and creative space. The other research trail follows the film editor’s work as a designer, from a perspective of knowledge, skills and intentions, ending with film viewers’ response to the completed work. The two trails meet and intervene where knowledge and skills are needed for the creative work as a film editor. Parts of the studies have been accounted for and presented on several occasions, and at different points in time (see List of Papers), but are here compiled per research project.

The first study is a theoretical analysis of film and TV production literature where established knowledge from the field is presented in writing by experienced and renowned professionals, analyzed and discussed from a design perspective. The study was carried out from 2012 to 2014, and was presented in a paper accepted by The Design Collection, that discusses moving image (and sound) work as audiovisual design. In this study, I undertook the analysis, and was supported in the writing process by Professor Yvonne Eriksson.

The second study carried out, began in Fall 2009, and was completed in Spring 2012: New design processes in the audiovisual industry, and considered how design within the Swedish film and TV industry had to adapt its processes according to the requirements of new technologies. A key result from this study was that designers’ space for creative work, inter alia, is constrained by the adjacent material factors: time, material and tools/skills. These results were published in the Journal of Integrated Design and Process Science. In my Licentiate thesis (Swenberg 2012), design processes with several design agents (within film and TV production) are thoroughly discussed in terms of how the creative space that each design agent possesses possibly interferes with the creative spaces of nearby agents. A prototype for handling unnecessary time-loss in these kinds of design processes was developed and evaluated against industrial representatives. The results were presented at the SDPS 2012 - The 17th international conference: The technology of complex systems, in Berlin, June 2012. This study was conducted in co-operation with my colleague, doctoral candidate, Per Erik Eriksson, with whom I equally shared research efforts, and
with support of our supervisors, Professors Yvonne Eriksson and Árni Sverrisson. This thesis further develops parts of the results of the study, which are discussed in my Licentiate thesis: processing of materials (in which perception plays a role) as part of what claims creative space for a designer.

The third study starts with a pilot-study, *Klippunkter, klippare och visuell perception*, carried out between 2009 – 2010, and presented at IV10 – 14th international conference information visualization, in London, July 2010. This paper presents a research design for studying film editors’ intentions with the edits created in films, in order to later compare these with viewers’ perceptual responses to the edits. I carried out this pilot-study on my own, with consultation support from my supervisor, Professor Yvonne Eriksson. This study was developed into the project that became the main study supporting this doctoral thesis. The full study was titled *Klippares visuella intention och tittares visuella perception*, and was carried out from 2012-2015, and presented in three papers. The first paper systematically describes and discusses the method, presented at JVMC4 – 4th international visual methods conference, in Brighton, September 2015, and further developed and published in the journal, *Visual Methodologies*, 2016. The second paper presents how a film editor involves perceptual considerations, and is published with the *International Journal of Visual Design*, 2017. The third paper accounts for viewers’ visual responses to film edits, and is under review by the journal, *Empirical Studies of the Arts*. This main study was carried out largely by my own efforts, assisted by my fellow doctoral student, Per Erik Eriksson, at Dalarna University, during the runs of the eye tracking experiment.
Delimitations

The concrete objects of study in this thesis derive from the field of film production. This field employs visual design, as well as design of sound in relation to visuals. The results from those studies presented and discussed in this thesis should primarily be considered as valid, or credible, within the contexts from which they stem.

The first object of study is documentary film editing, as an instance of film editing for continuity. Documentary film editing provides an interesting case for studying perception at the edit points, since it is harder to achieve continuity with such a film material, than with, for instance, fiction film material (cf. Nichols 2001, Hampe 2007, Rabiger 2014). These challenges are here assumed to be more evident than when fiction film shots are edited, since such shots are closely planned to be joined before shooting takes place. Editing is then usually not as constrained by image features that do not fit each other across an edit. Further, film editing has to consider the combination of sounds and moving images. It is here assumed that audio may also influence seeing, just as visuals may influence hearing (Matlin & Foley 1997, Anderson 2005). Thus, during editing, sound is regarded as a possible factor affecting a viewer’s visual perception. However, in the study of viewers’ perceptual responses to film edits, the second object of study, it is only visual attention that is measured, not visual effects on hearing. This is the primary reason for the use of parentheses when writing “(audio-)visual” throughout this thesis.
Theory and Related Research

Theory, I understand, as an overarching term for academic sets-of-concepts for the explanation or interpretation of data, sets that may be turned into models: “A collection of concepts, even a small collection, may be assembled in some logical fashion that then might represent a theory about the events that have been studied” [emphasis by Yin], (Yin 2011). Largely, the theory I use is directly derived from previous research. Therefore, I will approach theory and previous research together, highlighting the core concepts from the research I present. I will address three research areas, design, media production, and perception, from which I will summarize the concepts, and describe how I interpret and use them. Concepts that are usable as operationalized quantities to describe data will be addressed accordingly in a separate section.

The broader scope of theory I will consider departs from existing models and concepts in previous research. Each concept and model will first be illuminated from the perspective of its origin, and its relevance to this thesis will be discussed. Then, I will match models and concepts according to how they, in combination, can contribute with explanatory power, in relation to the purpose of this thesis and its research questions. I will here describe how the models are used and in what way.

Related Research and Core Concepts

In the following, previous research is presented from its respective point of departure, methods used, results, and contribution of theory. Core concepts that I re-use in one form or another will be given particular consideration. Three areas of research will be included in this chapter.

Design Research

This section will outline some design research of relevance to this thesis, with reference to the understanding of design, and how it relates to perception and attention: definitions of design, research on design professions, audiovisual design, as well as perception in design.

One notion of design is that designing should be considered as activities that shape materials in order to make things that are meaningful to users (Krippendorff 1995). Klaus Krippendorff, in his review of the design research
community’s understanding of what design is, suggests an axiom as a point of departure: “design concerns itself with the meaning artefacts can acquire by their users”. He emphasizes that “humans act not on the physical qualities of things but on what they come to mean to them” [Krippendorff’s emphasis], (1995 p.150). The meaning of a thing evolves during the course of someone interacting with it. The resulting theoretical stance is that: “Artefacts always afford many meanings. […]. [D]esigners can do no more than provide the affordances for users’ meaningful involvement” (Krippendorff 1995 p.152). How an audience can make meaning of a film or a TV show, as intended, will be discussed as an issue for those producing films and TV.

Another notion of design is Herbert Simon’s idea that design is “aimed at changing existing situations into preferred ones” (1996 p.111). Simon addresses issues regarding how artificially made phenomena can be studied from a scientific perspective. He uses existing examples from management, business, industrial systems, and engineering for his analysis. The results are theoretical constructs: design is, in Simon’s understanding, occupied with solving problems. Such problems should be elaborated and defined, and provided with desired functions or distinct goals, including criteria that alternative design solutions can be assessed against (Simon 1996). Simon admits, though, that artistic design might have to rely on continuous development of goals that are not fixed from the start. To the extent that film and TV production is a case for this kind of artistic design could lend itself to a discussion on problem-solving.

Some research of relevance to this thesis concerns how design processes, design issues and problems, and design work are understood. Much design process research is focused on understanding all parts of a process, its components, sub-processes, and activities (Eckert & Clarkson 2005, Poggenpohl & Sato 2009). Even as design fields, design contexts, and design purposes differ within the broad range of design research, it is common that the designer must master the design process, and even though “the problem of bounded rationality” prevails, must be able to make decisions and carry out activities accordingly (Horváth 2004). A core aspect of design, emphasized in design research, is the employment of design principles (Pahl, Beitz, Feldhusen, & Grote 2007, Blair-Éarly & Zender 2008, Poggenpohl 2009). Poggenpohl (2009) suggests that “meaning-making” and “problem-solving” should be understood as two such principles, and that their employment integrates them during a design process, starting by addressing a problem, and ending with considering how something becomes meaningful to its users.

One recurring design issue is how to reach, and not jeopardize, the involvement, emotion, and experience of the design, on behalf of the user (Lewalski 1988, Bloch 1995, Jordan 2002, Sato 2009). The user must be attracted to, and engaged by, the design (MacDonald 2001). At the end of co-creation, this means that a designer must pay considerable attention to aesthetic features, such as shape, size, color, texture, symbolic affordances, and more (Sato 2009). Not
least within audiovisual design, these aesthetic considerations are manifested in the designer’s treatment of materials, which in the audiovisual case, are digital materials (Paper B). Materials, digital and otherwise, must be treated aesthetically in such a way that they attract and engage users and viewers, making them become involved and affected, and giving them a palpable experience. This requirement must be observed when studying the profession of film editing.

Design Professions
There is an extensive body of academic research being carried out on design professions: how they are distinguished, how they come to accustom to an existing organization, how they are adopted by aspiring recruits, how they are taught and learned at educational institutions, among other things. Industrial design and engineering design have gained most attention, but some conducted research regards design work, and particular professions that have more of a “crafting” character, or appliance of design skills in specific contexts with demarcated design purposes. This research is relevant in regard to the study of film editing as an executed activity by a professional film editor.

When studying some trades, it is not immediately obvious how they are best categorized, since they include activities that are instances of crafting, while the outcome could be called either art or design. Kristina Niedderer and Katherine Townsend (2014) address the issue of understanding crafting as design (or art), and the kinds of knowledge embedded in such craft professions. Their approach is to find a research stance that can capture the nature of such knowledge, as they review research literature from the areas of arts and humanities (including design), in the UK, spanning the last two decades. The results suggest new methods for researching crafts that address the crafting activities, not only materials, technologies, contexts and practice models, but methods that are capable of capturing engagement aspects, such as intimacy and emotion. Theoretically, Niedderer and Townsend contribute by demonstrating how crafts can be understood as “to fluctuate on a continuum between art and design” (2014 p.626), where parameters of analysis are decisive for the outcome. Central concepts are experience and emotion, while they suggest several usable perspectives: aesthetics, expression, function, technology, skill, and quality, amongst others. Experience is regarded as “the actual observation of facts or events, considered as a source of knowledge” (Niedderer & Townsend 2014, with reference to OED, 2010), while emotion is considered as “the personal and individual response to an experience”. Film editing has been identified as existing somewhere in between craft and art (Orpen 2003), which confirms the relevance of such an angle in addressing the trade of film editing.

In her cognitive studies of creativity, Ingar Brinck (1999 & 2007) narrows down problem-solving further, using explicit aesthetic problem-solving arts and crafts as examples. She, also, regards artistic creativity as aesthetic problem-
solving with an undecided end, typically distributed and context dependent (Brinck 1999 p.34). Brinck’s analysis uses common examples while regarding existing theories of creativity, problem-solving, context-dependency, situated cognition, dynamic systems, and aesthetic experience. Her results are a theory that suggests understandings of creativity in context. Brinck considers artistic creativity as “an embodied, experience based craftsmanship” (Brinck 2007 p.422). This craftsmanship manifests itself, inter alia, in the individual agent’s relation to the material processed, as well as in his/her exploitation of some technical equipment used for tooling the material. I will re-use these notions for film and TV production, when I discuss tools and materials as digital, as used by individuals for aesthetic and communication purposes.

Another perspective for understanding professions is to capture how others than those exercising a profession comprehend that profession. Gillian Smith and Allan Whitfield (2005) studied how well design stakeholders recognized different design professions, from the point of intellectual categorization. Their results show that some design professions are well, or moderately, understood by the public (government agencies, funding bodies, industry, design commissioners), whereas others were rather poorly understood. There was a lack of distinguishable categories for sufficiently identifying several design professions. Some of these categories are what work a profession is engaged in, and Smith and Whitfield (2005) found that the public did not know what kind of work some design professions do. These results stem from an Australian large-scale investigation where numerous people from the design community, design education programs, as well as the public participated, conducted through “an occupational prestige survey”, focus groups, and a widely distributed questionnaire survey. Smith and Whitfield’s (2005) theoretical contribution is the use of prototype theory by which individuals can comprehend others as belonging to some sort of pre-established category, for instance, a design profession. Each such category is established by means of some topography, a set of characteristics that makes it possible to navigate within this category. It is the lack of such topography that thereby makes a design profession hard to identify. Film editing has existed for little more than a century, and is described as “the defining art of cinema” by Walter Murch (2006 p.x), that “invented itself in a cocoon of silence, and [has] continued that reticence […]”. How film editing can be scanned for its topography as a design profession will be presented in the Analysis chapter, for further discussion in the Discussion and Conclusion of this thesis.

How a professional designer becomes one, and continues to employ “design thinking” as a designer is studied by Robin Adams, Shanna Daly, Llewellyn Mann and Gloria Dall’Alba (2011). They analyze two phenomenographic studies of design practices through a framework for embodied understanding of practice, that interconnects knowing, acting, and being as a professional. This is what is understood as specific design knowledge and thinking, typical design actions, and what attributes are unique to being a
designer. The theoretical framework they use to interpret designers’ experiences of designing includes: “[…] (1) what professional designers come to understand about the nature of design and design problems (thinking), (2) how professional designers approach design (acting), and (3) how design professionals perceive design identity (being)” (Adams et al. 2011 p.594). Their results are theoretical constructions regarding kinds of experience of designing across different design professions that formulate different understandings that all include the three aspects of knowing, acting and being. Film editing is formulated as a profession by experienced and renowned film editors and scholars, which are statements lending themselves to similar interpretations (Crittenden 1995, Fairservice 2001, Orpen 2003, Pearlman 2009). Specifically, the use of perceptual thinking in film editing is identified as an issue for further empirical research (Smith 2005).

Audiovisual Design
Audio-visuality in design can be addressed either as occurrences of coupled audio and visual components in any design context, or as the specific design of audio-visual objects. In this thesis, the focus lies on this latter form of design as audiovisual design. Björn Bartholdy (2008) describes “audiovisual design” in the Design dictionary as integrating sound and moving images, either as film design, TV design, or as animation. In this understanding, many film and TV professions, primarily production designers and broadcast designers, are occupied with audiovisual design. This understanding of film and TV production as audiovisual design informs my previous research on such production processes (Swenberg 2012), and is argued for in Paper A.

Actual research on audiovisual design is rare, and some few examples I have come across regard either moving graphics (e.g. Pertíñez-López 2015), graphics in relation to moving images (e.g. Las-Casas 2013), or multimodal semiotics (Ponte & Niemeyer 2013). They all engage in analyzing the design, or composition, of graphical elements in relation to audio in existing audiovisual objects. Perspectives used are aesthetics for communication, or multimodal communication. Raquel Ponte and Lucy Niemeyer (2013) contribute to the theory by applying Lucia Santaella’s semiotic matrixes to film, which emphasizes audio-visual complexity. Jesús Pertíñez-López (2015) shows that aesthetic and artistic quality that is successful with the audience is repeated in later productions. Luiz Fernando Las-Casas (2013) investigates the principles applied in film typography and their effect on communication. The interaction of audiovisual components as semiotic features of communication, in need of designerly adherence (cf. Kress & Van Leeuwen 2001), will be regarded in the analysis of documentary film editing.

Perception in Design
How design practice has come to embrace the stance of regarding humans as acting in the world is discussed by Koskinen, Zimmerman, Binder, Redström,
and Wensveen (2011). In *Design research through practice: From the lab, field, and showroom*, they survey a volume of international research from the field of constructive design, and find that design practice has abandoned the position of viewing users as semantically interpreting designs. The fundamental shift, according to Koskinen and colleagues (2011), lies in the understanding of perception in design. The user’s semantic perception as a cognitive process of interpreting a product’s symbols and signs is contrasted with an approach built on Gibson’s direct perception (Koskinen et al. 2011, with reference to Ross et al. 2008). The direct perception approach sees humans as situated in the world and as action-oriented. In this understanding, designers have to adhere to the direct perception approach by letting the affordances of things be apparent for immediate user interaction. This action-oriented approach to design, based on direct perception, “can be turned into a design tool” (Koskinen et al. 2011 p.112). In my analysis, both aspects of design perception, semantical and action-oriented will be addressed for further discussion.

Péter Pál Boda and Yuan-Yi Fan (2013) present a study of interaction in digital storytelling for experience sharing, where multisensory perception is a crucial part of the framing. They explore a multimodal platform/application, called *ADEO*, by designing and implementing it, as well as testing it for user experiences. As a result, they manage to successfully integrate speech recognition, 3D gesture recognition, physical activity recognition, and proximity sensing as inputs to the system (Boda & Fan 2013). Theoretically, Boda and Fan (2013) contribute by providing conceptualizations for how multisensory perception contributes to experience sharing. Multisensory perception will be addressed in my analysis and discussion.
Academic Media Production Literature

Just as design research only to a limited extent touches upon audiovisual matters, media literature rarely brings about a design perspective. Media production is discussed in research literature as well as in other academic writings, both of which will be presented here. Media concepts, such as genre and industry divisions, will be addressed, as well as film and TV production contexts as cultures of, and conditions for, production. The coupling of media production to the area of design will be in focus. Particular interest is devoted to film editing.

First writings on film production occur in the early twentieth century. Hugo Münsterberg describes how film is constructed, from a psychological perspective, and some of its particularities that otherwise would have remained hidden to those outside film production (2002[1916]). He elucidates how arrangements and details considered in production are deliberately intended to attract the spectator’s attention. In reference to editing, Münsterberg states that “the mere formal arrangement of succeeding pictures may keep our attention in control” (2002[1916] p.84). This “controlling” function of film editing is part of the topic of this thesis.

For a considerable length of time, there has been an identified need for research into film and TV production in order to theoretically understand the knowledge and processes of the field (Finn 1953). The framework of such research is presented as being able to expand knowledge by bridging and integrating theory with film and TV production practice, as well as developing and strengthening professions within that field (Finn 1953). How films are designed has been another issue, either from an economic viewpoint (Wagner 1955), where the conceptualization of film aesthetics is regarded from the film’s potential for returning investment, or from the marketing perspective of addressing demarcated audiences (Levonian 1960), where the concern is the relation between a film’s design and an audience’s appreciation of the film. More extensive discussion on film and TV production as audiovisual design is carried out by Wim Westera (1995& 1999), as he discusses particular crafting, efforts and sub-processes that concern the compilation and completing of audiovisual objects in film and TV production processes. Westera’s framing of film and TV production as audiovisual design is concerned with how the development of an audiovisual object can be addressed as a design problem with many design decisions to be made during the production process (Westera 1995). This thesis attempts to bridge the gap between perceptual understanding of designing and its practice, where film editing is one such instance.

Yet a coupling of media production to design is Lars Nyre’s (2014) suggestion that new forms of media communication should be developed according to information design principles. The exploitation of new technology and communicational means needs a more thorough approach when their poten-
tial capabilities and uses are examined, for reaching new purposes and applications (Nyre 2014). This design approach to media production is consistent with the approach of this thesis.

Another approach to researching the film and TV industry is taken up by John Caldwell (2008), as he applies a culture-critical perspective when examining Hollywood for empirical data. Caldwell uses mixed methods: interviews, observations, audiovisual objects, and gathering of industry-generated documents, which he analyzes for individuals’ self-conceptions and power distribution within the industry. The output regards how economic and social order within film and TV production in Hollywood is influenced by industrial structures, technological factors, organization, and social relations. One of Caldwell’s core concepts is the “critical intelligence” that each individual has to uphold in order to understand situations from a production and organizational perspective, as well as what is required by them to meet the demands of a production, and to persist in the industry. Another result is the identification of the film and TV industries as merging, where the same people and firms are engaged in film production projects, just as in TV production projects (Caldwell 2008).

**Documentary Film Material and Editing**

One aspect of this merging is addressed by Per Erik Eriksson (2013) in his Licentiate Thesis, *Videography as production nexus: digital film cameras, media management and the distribution of creativity in TV and film production*. When film and TV production completed their entrance into the digital era by adopting high resolution digital video cameras, the roles of the image creators (cinematographers, videographers, and camera crew) changed; they had to share their creative space with other designers further down the production line, simultaneously as their roles expanded as gatekeepers and providers of digital material proper for the rest of the workflow (Eriksson 2013). Proper digital material is crucial in maintaining creative space for all designers throughout the workflow. Eriksson has addressed the issues of digital video material and its qualities in two reception studies with screenings and questionnaires. In one of the studies, he observes that young viewers tend to prefer digital images that have textures resembling older technologies, over more high-resolution digital video images (Eriksson P.E. & Eriksson Y. 2015). In the other study, Eriksson observes that digital video image materiality created by convergence cameras lessens documentary film viewers inclination to ascribe credibility to the image due to its (high) aesthetic appearance, as resembling fiction film images (Eriksson 2012). These results are indications of viewers’ sensibility to the material quality of moving images, and that their reception affects their understanding of the images, and by inference, the films in which the images occur. These conclusions emphasize the importance of adherence to such qualities when designing moving images.
Documentary film professor, Michael Rabiger, in his extensive work that describes and clarifies documentary film production, *Directing the documentary* (2009), addresses, amongst other things, the specific complexities of editing a documentary film material. He articulates the editing process as a second chance to direct the documentary film:

Seldom does a film follow the kind of scripting through time that determines a fictional story. [...] Editing becomes like making a picture from found objects. Individually and in juxtaposition, they suggest what pictures are possible, and you start experimenting. Such searching and experimenting in the cutting room usually arises because shooting has taken its own pathways and turned out differently from what you expected.  

(Michael Rabiger 2009 p.206)

In documentary film editing, the editor cannot depend on a previously established script that the shooting has followed, since the shots are usually improvised, and not planned to be joined smoothly, as they are when a fiction film is produced. Instead, the editor must create connections between shots in documentary film editing. Moreover, the story might have to be created or re-modelled during editing, from a stock of audiovisual material that is diverse: first structured, then refined to achieve a flow. Rabiger identifies these aspects as two main problems that must be solved during any documentary editing, “Editing a documentary is similar no matter whether the film is short or long” (Rabiger 2009 p.207). These are reasons for choosing a documentary film sequences to be edited in one of the studies underpinning this thesis.

In an analysis of a number of well-known twentieth century films and literature on film editing, Valerie Orpen localizes film editing exactly on “the line between art and craft” (Orpen 2003 p.1). She points out that film editing is always context dependent, contingent on footage, genre, acting, sound, and more. Therefore, editing is also complex to analyze. Most evident, Orpen claims, is the interdependency of shooting and editing, where shots are created with editing in mind, and editing is dependent on the existing footage. The contribution of editing is either to support the narration and comprehension of a film through connecting, invisible, smooth edits that escapes the viewer’s attention, or to be expressive and therefore visible, obtrusive, and attracting attention. Orpen’s conclusion is therefore that it is impossible to isolate editing to the analysis of “the cut per se” [emphasis by Orpen], (2003 p.3), where one shot meets the next. This thesis attempts to do just that, namely, to examine the encounter of a specific frame of one shot, with another frame from another shot.

Directors Karel Reisz and Gavin Millar (1968), in *The technique of film editing*, start out by establishing that the practical achievements leading up to continuity editing, with examples derived from the history of film, came in direct succession to film editing itself. Film editors, they claim, have since made efforts to refine the smoothness of continuity:
Harsh, noticeable cuts tend to draw attention to technique and therefore tend to destroy the spectator’s illusion of seeing a continuous stream of action. Constructing a smoothly flowing continuity has, indeed, become one of the modern editor’s main preoccupations. (Karel Reisz & Gavin Millar 1968 p.48)

Reisz and Millar describe how in a scripted film, scriptwriter and director should take responsibility for how continuity is achieved in the editing. Conversely, in a documentary film, the creation of the film depends on what “can be produced by the factor of editing” (Reisz & Millar 1968 p.123). Smooth continuity is then subordinate, they state, and harder to achieve.

Later research on film editing by Karen Pearlman (2009), Cutting rhythms: Shaping the film edit, focuses on the elusive concept of “rhythm” in editing, which every film editor seems to return to when making a statement on film editing. Pearlman seeks support in theories of bodily actions, and compares editing rhythm to rhythms in the world and to bodily rhythms. From that, she claims, film editors learn to develop sensitivity to naturally occurring rhythms that they employ when they shape the rhythm of film events. Other aspects of shaping film edits, that Pearlman discusses, are collision and linkage between shots, two concepts she adopts from early film thinkers Sergei Eisenstein and Vsevolod Pudovkin, respectively. Pearlman’s understanding of collision is not equivalent to discontinuity, but rather refers to graphical properties of the image, as well as to movement. Similarly, her use of linkage seems very close to what others call continuity, especially when linkage is smooth, but again the concept primarily regards movement and image form (Pearlman 2009). Any film edit, she states, will end up somewhere on a scale between collision and linkage. I will use all four concepts, collision and linkage, continuity and discontinuity, and will then use this interpretation of Pearlman. I will also discuss rhythm in film events, and then in relation to event segmentation as a perceptual phenomenon.

Digital Material and Creative Space

Achieving collision, linkage, rhythm or other aesthetical qualities in film editing depends on the conditions for creativity. My own research on film and TV production processes departs from the point of technological shifts, and their influences on such processes (Eriksson P.E., Swenberg, Eriksson Y., & Sverrisson 2010). The analysis of interviews with industrial managers, and a post hoc analysis of complex production reveal that film and TV production processes are sensitive, or not, to whether each audiovisual design agent in a workflow of many agents has immediate access to the proper digital material for tooling (Swenberg 2012). The theoretical contribution is a three-dimensional model in which the material is linked with the tools in use, combined with user skills, and with time available for designing. The model displays a creative space for each design agent, where the creative space of one agent is
part of the context of the surrounding agents, and thus influences their respective creative space. The processing of the audiovisual material is the essential design activity, and how to address issues of a film editor’s processing of digital materials will be presented in the Analysis chapter, below. Further, digital audiovisual material is long since recognized as not being a straightforward matter in itself (Manovich 2001 & 2012), since there is a great variety of digital configurations (codecs) to employ, bringing new potential and conditioning the material from expressive points of view.

Perception Research

Perception, as a tool in design creativity, will now be elucidated in regard to previous research. As already outlined in the Introduction chapter, perception is understood and approached differently in different academic traditions. I will approach perception from two directions, namely, from a cognitive, and cultural viewpoint. In order to acknowledge context, I am dependent on cultural and societal understanding. Simultaneously, when dealing with practical details in regard to techniques, technologies, systems, and processes, cognition has become a recurring perspective.

From a cultural studies perspective, perception is a means by which a being selects informative input, not just of a sensory kind. Perception is the processing moment where sensory or intellectual input is merged by the mind with previous knowledge: memories, emotions, preferences, interests, learnt matter, and more (cf. Fry 1981, Hamlyn 1996, Crary 1999). The intention of the being is what directs perception to whatever seems to be of interest. What actually is perceived is therefore intrinsic and unique to every individual. Perceiving is then a part of his/her process of understanding the surrounding world and each situation therein.

Understanding is also used, coupled to the notion of reception, which renders a similar explanation. Therefore, I am not convinced that there is any significant difference between the concepts of perception and reception in this tradition. From a cultural studies perspective, I will stress the importance of intention for design perception, and consider reception as the process of decoding and integrating new inputs, sensory or intellectual. Accordingly, only the decoding aspect will be actively used in my study, while the notion of reception will be disregarded.

However, it is evident that some perceptual functions claim privilege to the mind’s attention capacity, completely separate from the influence of intention: for instance, immediate reactions to sudden noise or flashes. In cognitive science, perception is one of the physiological systems that a body has for taking in information to the brain about the local environment and surroundings (cf. Pylyshyn 2003, Goldstone et al. 2015). These information inputs function regardless of what occupies the mind at that moment, and do so involuntarily, and independent of any kind of thinking. The brain processes
sensory input all the time, though not consciously or attended to. I will draw on these immediate aspects of perception when discussing perception in design.

**Gibsonian Direct Perception**

When immediate perception is discussed within design research, the work of James Gibson (2015[1979]) is often referred to. Gibson’s suggestion, that visual perception is “direct”, means that the visual processing system on very short notice recognizes properties from the surrounding environment, and identifies its “affordances”, consciously reflected on by the individual, or attention undisturbed. His understanding of perception is that it is “ecological” in the sense that it belongs to a living embodied being with many senses, and in constant active interplay with its environment. This active manner includes moving in space, viewing binocularly, keeping balance, being supported by the ground, and more. During that interplay with the environment, a being always has a “point of observation”, Gibson states (2015[1979] p.59), which changes with movement and new positions, and from which the being can perceive what properties of the surroundings are unchanging, or “invariant”. These invariant properties, in combination, constitute the affordances that the environment and its objects offer us. With affordance, Gibson means something more than a property, that is to say, a capacity-possible-to-employ (2015[1979]). His theory of ecologically conditioned direct perception of object’s affordances explains how we can perceptually interplay with our environment (Turvey et al. 1981).

Affordances should be understood as dispositions, if we concur with Andrea Scarantino (2002), since they are of many different kinds, and depend on our deeds and abilities in order to be perceived. An object can have a certain disposition that might be employed by a being, but it depends on the being’s capabilities whether or not a disposition is perceived as an affordance. Harry Heft (1989) argues for the possibility of extending the concept of affordances to instances of culturally-derived functional meaning. For example, when seeing a switched-off TV screen, one cannot see, from its direct appearance, that it is capable of holding a lit picture, but most sighted individuals that are part of an industrialized society perceive that affordance as a function, which is learned within our culture.

From here, the leap to the application of the concept of affordance within design is not very great. Donald Norman (1998) uses the notion of affordance to indicate how designers ought to consider how users perceive items, and that functions and usages should be directly evident to them. I will return to the concept of affordance when discussing film editing.

**Eye Movements, Perception, and Attention**

How the affordances of things are perceived is directly linked with human attention. Therefore, I will now turn to studies of how the eye responds to visual stimuli, conducted within the field of perception research. Early studies
tried to figure out the correspondence between visual stimuli and eye movements (e.g. Stratton 1906). Guy Buswell (1935) is accredited with the first major eye-tracking study regarding pictures. Buswell found that the eye stays at fixed positions (fixations) and makes short, quick movements (short saccades) between fixation at some part of a picture that, for some reason is more interesting, followed by long saccades (long, quick movements) to some other part of the picture of interest. This way he also found gaze patterns that seemed more “local” to limited parts of a picture, and “global” patterns that seemed to correspond with making an overview of the picture as a whole (Buswell 1935). Another aspect of picture viewing that Buswell noted was that tasks influenced viewers’ viewings (1935). This aspect was thoroughly examined in an extensive eye-tracking study by Alfred Yarbus (1967), who found that there are many ways – in terms of gaze patterns – to look at a picture, and that each task given to a viewer renders a unique gaze pattern for the same picture. But, generally, the gaze is drawn to representations of faces, human bodies and animals (Yarbus 1967).

From these results, the relation between the human gaze and attention needs further illumination. Victor Lamme (2003), in considering the difference between awareness and attention, identifies attention as the condition under which “some sensory inputs are processed faster or deeper than others, and thus become more readily available for action, memory, or thought”. I will fall back on this Lammeian formulation of attention. There is also growing consensus on a strong co-location of attention and eye fixations when viewers look at a visual stimulus, as Jason McCarley and Arthur Kramer (2007) found when surveying the field.

Attention, though, is not a simple phenomenon, according to Howard Egeth and Steven Yantis’ (1997) review of psychological research on attention. They identify three major aspects of human visual attention: (1) goal-directed (top-down) versus stimulus-driven (bottom-up) processes; (2) the representational basis (locations or objects) for visual selection; and (3) the time course of sequential attention to stimuli (Egeth & Yantis 1997). In addition to that, Yair Pinto et al. (2013) have claimed that bottom-up and top-down attention are two independently functioning systems. This suggestion entails that stimulus-driven attentional capture is not contingent on goal-directed steering of attention, i.e. intention. Pinto et al.’s result means that the two processes, top-down and bottom-up, compete for access to attention (2013). In filmmaking, mastering this competition between attentional systems becomes a delicate matter (cf. Batty, Perkins, & Sita 2015).

**Perception and Attention in Film Viewing**

A recent attention study on film viewing (Cohen, Shavalian, & Rube 2015) revealed that the cut-up of a film, out of story order, affected the viewers’ attention significantly. Viewers were less able to remember pre-screening instructions while watching an engaging and continuously running film,
whereas the cut-up version did not attract attention to the same degree, and viewers were able to remember the instructions (Cohen et al. 2015). All these major aspects of human attention will re-emerge during my discussion of results below.

Perception psychologists have directed attention to film for some time. A major investigation was carried out by Julian Hochberg (1986), regarding how some constitutional principles of moving images appeal to our visual perception. There, he deciphers how screen construction, camera movements, optics, edits, and more, relate to perception (Hochberg 1986).

More recently, interest in where an audience actually looks while watching a film has increased, with the availability of tractable and modestly priced eye-tracking equipment. Robert Goldstein, Russel Woods, and Eli Peli (2007) studied whether people look at somewhat similar areas of the film image. Their interest was in finding ways to improve viewing for visually impaired. They found that for circa half the running time of a film most participants looked at an area smaller than 12% of the film scene (Goldstein et al. 2007).

Paul Marchant et al. (2009) examined over 400 viewers watching Hitchcock’s film *Vertigo*, with the purpose of evaluating the influence of the production techniques used in creating the film, in relation to the viewers’ gaze. They found that cinematic efforts to guide the spectators, on the whole, worked as estimated by the film creators (Marchant et al. 2009). Yet, eye tracking can be employed not only in the study of how the many view film, with the purpose of controlling attention (Brown 2015), but can open up to the study of a variety of individuated ways to watch film (Robinson, Stadler, & Rassell 2015), as well as being combined with other, e.g. ethnographic, approaches in order to understand film viewers’ world comprehension (Dyer & Pink 2015).

The effect of film creators’ ways of directing the audience has also been studied from the perspective of interpreting filmic events (Cutting et al. 2012). James Cutting, Kaitlin Brunick, and Ayse Candan (2012) let spectators indicate where new events of film sub-sequences started, and analyzed the relation between indicated event breaks and the film’s visual construction (shot scale, motion, edits, color shifts, and more), and found strong correlations between events and film construction. Cutting (2014) concluded that film viewers identify event shifts better the more these shifts are constructed by means of film discontinuity elements. Film editing can support and take advantage of event segmentation (Magliano & Zacks 2011). Continuity editing is closely related to the perception of events’ starting and ending, according to Joseph Magliano and Jeffery Zacks (2011). Further, they found that edits that do not follow events are reacted to by viewers as if action is not continued. Whether event segmentation is concretely applied in actual film editing is not addressed, though. In my study of a film editor’s intentions and viewers’ perception, I have regarded event segmentation in my analysis of the film editor’s work, as well as when regarding film viewers’ eye data at film edit points.
Perception, Attention, and Film Edits

In a series of studies on edits, perception psychologists have paid specific interest to cuts in moving images. Julian Hochberg and Virginia Brooks (1978) study “visual momentum” as a film editing means of drawing viewers’ attention across film edits. They suggest that a way to achieve such visual momentum is to pose “visual queries” – attractions for perception to follow as leads to new discoveries – before an edit point, that is then answered in the next shot (Hochberg & Brooks 1978). Hochberg and Brooks (1978) do not address actual editing processes on these matters, as I have done. My data revealed other actual exploitations of visual queries that a film editor takes on, which rendered additions to Hochberg and Brooks’ theory, the concept of “visual query exploitation” (Paper F).

Géry d’Ydewalle and Mark Vanderbeeken (1990) contested continuity editing rules by violating them, and gathering viewer eye data. They found that perceptual reactions were significantly higher when the rules were violated (d’Ydewalle & Vanderbeeken 1990). Further, d’Ydewalle & Vanderbeeken tested viewers’ response times after violated and non-violated edits, as an indication of cognitive load, and found that violating editing rules contributed to an increase of cognitive load in the form of slower responses (1990).

In a follow-up study, Géry d’Ydewalle, Geert Desmet and Johan Van Rensbergen (1998) could confirm these results and establish a time interval where the amount of eye movements after a film edit is expected to increase if the edit is harsh for perception. Filip Gerveys and Géry d’Ydewalle (2007) noticed an effect on viewers’ responses to a cut where a small movement, unrelated to the narration, occurred just before the cut. However, they did not investigate this phenomenon further. Conversely, such effects are provoked in my study of documentary film editing and viewers’ responses, whereby I rely on the time intervals established by d’Ydewalle (1990 & 1998) for my results.

Seth Geiger and Byron Reeves (1993) accomplished a related study on cuts in TV programming. They tested the impact on attention from cuts between scenes with related content, and scenes with unrelated content, and found that related content was easier for viewers’ attention to follow, than the unrelated content. They also found a stress on cognitive load for the unrelated sequences (Geiger & Reeves 1993).

Since continuity and discontinuity in film editing is of high relevance for this thesis, several studies on that topic will now be summarized. Tim J. Smith (2005), in his doctoral thesis, investigates the conditions for what constitutes continuity editing, from an attentional perspective. He arrives at three dimensions of continuity that must be considered in order to achieve and maintain continuity: time continuity, space continuity, and object continuity. Smith categorizes edits according to adherence, or otherwise, to each of these three dimensions of continuity. Object continuity has privilege, according to
Smith, and viewer expectations on spatio-temporality can support expectations on “existence constancy” of objects. Meeting these expectations in the editing supports the experience of continuity for viewers (Smith 2005 p.211).

The task for the film editor who wants to achieve continuity is thus to manipulate attention by accommodating each film edit, so as to reduce perceptual disturbances. These results indicate that the effects of discontinuity are: attention to the film edit, awareness of the editing, and inferior capacity of the viewer to understand and remember the content of the story (Smith 2005). Smith suggests that we understand the purpose of continuity editing as evading such effects. He also praises film editors for their much-developed sensitivity to visual perception, but does not conduct any studies on film editors per se.

As a further study, Tim Smith and John Henderson (2008) tested the continuity editing rules established by film editors (cf. Reisz & Millar 1968, Fair-service 2001, Berliner & Cohen 2011) on viewers, while eye tracking them. Between a third and a fourth of all continuity edits passed unnoticed by the viewers, whereas nine tenth of discontinuity edits were noticed (Smith & Henderson 2008). Thereafter, Smith and Henderson, with Parag Mital and Robin Hill (2011), also have eye-tracked video viewers watching natural scenes, and could establish that motion attracts viewers’ gazes to the extent that all participants looked at the same point of movement at the same time.

An attempt to validate the appreciation of continuity editing with viewers, by Heiko Hecht and Hermann Kalkofen (2009), was designed as letting participants action-match edits themselves, to the point where they were content with their editing. The result of this study was that the participants preferred edits with time-gaps as continuity edits (Hecht & Kalkofen 2009). These results were challenged by Arthur Shimamura, Brendan Cohn-Sheehy, and Thomas Shimamura (2014), when they tested film edits on viewers, using action-match edits with time-gaps, real time, or time-overlaps. They found their participants preferred time-overlap edits, contrary to what Hecht and Kalkofen (2009) had found (Shimamura et al. 2014).

Thereafter, Shimamura, Cohn-Sheehy, Pogue, and Shimamura (2015) have tested the effect of film edits on viewers’ attention. Participants watching feature film sequences were clearly inattentive the moment directly after edits (Shimamura et al. 2015). Further, they found that removing the sound-track from the film sequences significantly improved attention after edits. Their conclusion from these results is that the multimodal engagement in the narrative is strong on cognitive processes (Shimamura et al. 2015), a conclusion that aligns with other research on perception and cognition (McCarley & Kramer 2007).

This attention-seizing engagement in the narrative, through continuity editing, is addressed in a theoretical study by Tod Berliner and Dale Cohen (2011). They suggest a model for understanding how “classical” editing can create the illusion of continuity on behalf of the audience. In their model,
“imperfect and disjointed visual information”, as presented by continuity editing, is accepted by the human mind because this is similar to how we assimilate visual information in our natural environment. Humans disregard gaps and discontinuities in creating a mental model of the world, and the mind replaces real world impressions with filmic impressions and joins them in a parallel fashion (Berliner & Cohen 2011).

When studying eye movement strategies during naturalistic scene viewing of videos, Helena Wang et al. (2012) tested participants by scrambling video scenes in random order, and by randomized duration times. They found that viewers need half a second or more to re-establish a focus for attention when it was lost due to a cut (Wang et al., 2012). This will be considered as a time interval for lost attention after edits, when applicable.

None of the above studies attempts to approach actual film editors as they edit film, to investigate how perception is actually regarded during the handling of audiovisual material. This is what I have set out to do in this thesis, namely, to find the actual handling of perceptual matters during the processing of film material.
Applied Sets of Concepts

In this section, the concepts described and defined above will be connected into theoretical models, in order to contextualize the use of perception into the process of film editing. Some models describe existing configurations of concepts and theories, as I understand them. I add to some of those models either by connecting presented concepts or by suggesting new additional concepts; and some models are new constructs for application. I will use these models in my analysis and discussion, as I add to the transition from unarticulated to overt understanding of film editing as a profession, how it is design work, and the role of perception in film editing. I will also expand on my interpretations of concepts and how I couple them, as well as on how I use them. Some alternative terms and concepts used by others will be discussed.

Film Editing as Craft or Design Profession

It is not always easy to distinguish whether aesthetic crafting should be understood as art or design. The same applies to the outcome of such activities. Brinck seems to include both concepts under the larger concept of “artistic creativity” concerned with “aesthetic problem-solving” (1999 & 2007). As a polarized model with extreme end positions, pointing out directions for theoretical distinguishing, I thus use the idea of specific contexts with specific purposes, accompanied by appointed methods and defined principles to be significant characteristics of design. In this idealistic model, art is conceptualized as contrary to design, not accepting any discriminating boundaries, but adhering to the idea of “l’art pour l’art” (see Figure 1). Crafts are found on a sliding scale in between art and design, and may share properties from both, an understanding supported by Niedderer and Townsend (2014). This is also how Orpen (2003) understands film editing. In addition, perhaps most actual art and design activities would be found somewhere on that scale, as well.

![Figure 1](image.png)

Figure 1. Crafts are found somewhere in between art and design, and may share properties and characteristics from both.

The need for studying design professions so that they become distinguishable is advocated by Smith and Whitfield (2005). We need to understand what
characterizes different professions in order to communicate notions about design work in society. The idea of studying crafts from a research stance that can capture the kind of knowledge embedded in crafts, must, according to

**Figure 2.** A mapping of how a crafts research agenda, with its established main categories, should include crafting activities, and sub-categories, that aims to understand craft knowledge, according to Niedderer and Townsend (2014).

**Figure 3.** Film editing research mapped as addressing craft knowledge through crafting activities, and sub-categories, expanded from Niedderer and Townsend (2014).
Niedderer and Townsend (2014), address the crafting activities, since part of the knowledge is embodied in those activities. Artistic creative knowledge, as embodied and experience-based, is a view also held by Brinck (2007). Crafting activities are then found besides materials, technologies, contexts and practice models on the research agenda (see Figure 2). Engagement is the key category, which has sub-categories as, for instance, intimacy, emotion, and experience (Niedderer & Townsend 2014). Experience is open to scrutiny for objective evidence, e.g. perceptual responses, whereas emotion is a response internal to the individual, and thus dependent on testimony. Such a research approach to crafting activities lends itself to the use of several possible perspectives (see Figure 2).

When studying film editing, the active use of perception could be a type of experience to scrutinize, which I have done (Paper F), and the appraisal of rhythm could be a type of emotional response to try to understand, as Pearlman (2009) has addressed. In my study of documentary film editing, I use the perspectives of aesthetics, expression and function (see Figure 3). How editing principles are applied can be addressed either from an aesthetic perspective, following genre conventions or narrative style, or from an expressive perspective, as a means for artistic creativity and invention, or from a narration functional perspective of comprehension. The same perspectives apply when film editing concerns semiotic, multimodal aspects of adding or removing specific audiovisual details.

In addition, the perspectives of expression, technology and skill, suggested by Niedderer and Townsend (2014), coincide with two of the dimensions suggested to clarify a designer’s creativespace (Paper B), (see Figure 31, Results chapter). The understanding of a designer’s creative space expands on Brinck’s (2007) proposal that an artistic agent, for his/her creative purposes, is restricted by the material and the tools used. Tools, technology, skills, material and time are factors in the creative space model (see Figure 31).

By recognizing and understanding what characterizes each design profession, these professions will be recognizable (Smith & Whitfield 2005). The characteristic features of each profession will form that profession’s topography, by which it can be recognized and navigated (see Figure 4). These characteristics are informed by what kind of work is accomplished within that profession. From this metaphor, I will suggest some characteristics for film editing, recognized in earlier film editing literature, as well as found in the current research, in order to suggest a topography for public recognition.
Figure 4. A topography of a design profession that accomplishes A, B, C, D, and E categories of work, and their respective characteristics (Design Profession Topography illustration built on Smith & Whitfield 2005).

If the topography of a design profession is meant to support how it is recognized from the exterior, it might not totally reveal how the profession is understood from within. Adams et al. (2011) try to address such an understanding, mainly by the categories of knowing, acting and being, to see what is specific to a profession (see Figure 5). Designers’ knowing is here revealed in their thinking about design problems, their acting lies in their approach to designing, and their being resides in what they express about design identity. This kind of self-understanding concerning film editors already exists in the film editing literature (Crittenden 1995, Billinge 2017, Orpen 2003, Pearlman 2009). The coupling of knowing and acting in regard to perception in film editing is here of interest to me.

Figure 5. Knowing, acting, and being as interrelated categories of professional self-understanding (from Adams et al. 2011).
Film Editing as Audiovisual (Multimodal) Design

The notion of *design* has been subjected to many attempts of definition. I find Simon’s (1996) suggestion of design as problem-solving applicable to audiovisual design, just as Krippendorff’s proposal of design as meaning-making is also applicable. Design integration experience acknowledges the need for problem-solving thinking in the early stages of design, whereas meaning-making thinking seems to dominate the later stages (Poggenpohl 2009). From my understanding of audiovisual design, I consider problem-solving and meaning-making as mutually integrated perspectives, where the starting-point and core is aesthetic problem-solving (cf. Brinck 1999 & 2007), but the completion and inevitable outcome is the meaning made, and I will therefore use such an understanding of the design concept (see Figure 6). Nevertheless, before solving any problem, one must fully understand the problem (Schön 1983). Simon’s (1996) observation, that the set goals required for problem-solving might not be fixed from the start when it comes to artistic design is supported by Brinck (1999), as she discusses aesthetic problem-solving in arts and crafts. Audiovisual design embraces that openness, so the model should be interpreted accordingly.

*Figure 6. Problem-solving to Meaning-making model: Design understood as starting with a core of problem-solving (including problem-framing), but moving towards its completion where meaning-making is the thing.*

Regarding audio-visuality in design, the multimodal semiotic perspective can contribute significantly in illuminating how a complex audiovisual object communicates (Kress & Van Leeuwen 2001). This is a different kind of regard of semiotics to the one Koskinen et al. (2011) reject, since they consider the construction of meaning by ascription. The multimodal semiotic aspects must be touched upon in this thesis, since film editing clearly addresses semiotic aspects in some instances. However, the observation will be very brief and very basic. The point is that the moving image in itself provides representations for semiotic interpretation (B), as does the sound track (A), whereas the
combination of the two resources provides a common compound representation for interpretation (C), beyond the sum of individual meanings of sound and image (see Figure 7). Any change in A or B cannot ignore a possible change that also occurs in C. Audiovisual design must attend to this semiotic complexity (Ponte & Niemeyer 2013).

Figure 7. Audiovisual semiotics includes sound semiotics (A), image semiotics (B), as well as the multimodal semiotic merge of A and B in C, as something more than A+B. Changes in A or B should take a possible subsequent change in C into account.

When Westera (1995) addresses film and TV production as audiovisual design, the idea is that choices made, as well as stances and positions taken during the production process and its sub-processes, add up to a design problem, requiring design decisions, during the compilation of the audiovisual object.

Figure 8. Model of Westera’s (1995) conceptualization: the compilation of a film or a TV show as an audiovisual object can be broken up into a number of production efforts achieved by people from different crafts during a number of sub-processes. Taken together, these people, their efforts, and their processes deal with a common general design problem aimed at compiling the audiovisual object into a coherent one. This design problem requires a number of design decisions.

It is the craftspeople involved who face those issues, problems and decisions (see Figure 8). These are also the same people, and the same processes, that produce the audiovisual object, for that reason, we can consider the design
process and the production process as a merged one (Paper A, Kress & Van Leeuwen 2001).

The concept of a creative space that each designer inhabits is also derived from the film and TV production context (Paper B). A model of a designer’s creative space will be presented more thoroughly in the Results chapter, below. Here, I shall comment on Eriksson’s (2013) observation that a proper digital material must be provided and accounted for in a film or TV production process, since the look of a material is received by the audience, and interpreted as a layer of the message communicated through the audiovisual object (see Figure 9). Some viewers prefer old-technology-imitating image looks (Eriksson P.E. & Eriksson Y. 2015), whereas others find convergence camera images non-credible as documents (Eriksson 2013).

![Diagram](image)

**Figure 9.** Moving images material quality matters to viewers’ preference and ascription of credibility (Eriksson 2013, Eriksson P.E. & Eriksson Y. 2015). Material quality also matters to the possibility for designers to process the image and thus express things (Manovich 2001 & 2012, Swenberg 2012).

The different conditions between fiction and documentary film production are part of the framework for the research underpinning this thesis. From Rabiger (2009), and Reisz and Millar (1968), we can summarize the core differences as depending on planning and control (see Figure 10), where the fiction film uses a set script from which the whole production is planned: acting, design, shooting, editing, sound and postproduction. Shooting a script includes a script-breakdown, where each shot is carefully planned to match the script and the shots next to the current. None of these planning and control mechanisms prevails during documentary film production, consequently, the editor is faced with a lot more challenging story-making and shots to match.
In distinguishing how to conceptualize film editing, I will turn to Orpen (2003), Reisz and Millar (1968), and Pearlman (2009), (see Figure 11). According to Orpen (2003), editing is found somewhere between art and craft. I interpret the artistic aspects of editing to be more related to the capacity of expressiveness, whereas the crafting aspects are more associated with narration and comprehension. Film editing is dependent on acting, sound, and genre, according to Orpen (2003). She also understands editing as mutually dependent on the footage. I would substitute “acting” for “appearing” in documentary film editing. Further, I link sound and footage to semiotics through the sound and image details that an editor can choose to use or remove while assembling the film (see Figure 11).

Expressiveness and narration/comprehension, respectively, are supported through different core working principles in film editing. Using three different perspectives on film editing as a craft (cf. Niedderer & Townsend 2014) (cf. Figure 11), experience concerns the perception of film edits, where continuity is found to support narration and comprehension (Smith & Henderson 2008, Berliner & Cohen 2011), while discontinuity relates to expressiveness (Orpen 2003). Emotion concerns the inner response from viewing the edits. Expressive edits can be articulated as harsh; or smooth, if they seamlessly support narration, and do not disturb comprehension (Reisz & Millar 1968). From the material viewpoint, graphic and motional collision between shots relates to the expressive, while image graphical and motional linkage contribute to narration and comprehension (Pearlman 2009). Linkage and collision, in turn, also create different senses of rhythm (Pearlman 2009), (see Figure 11). Even though there is a relation between discontinuity and collision, and continuity and linkage, respectively, these relations will not be further expanded on. Collision and linkage will not be further used either, since the material aspects these concepts address, graphic form and motion, will not be regarded from that respect, but will instead be discussed from the experience...
perspective. Any differences between the handling of digital video and celluloid film when editing is beyond the scope of this thesis.

Figure 11. Model of film editing, its core working principles (cf. Reisz & Millar 1968, Orpen 2003, and Pearlman 2009). The perspectives of experience, emotion and materiality, respectively, provide principles that support either expressiveness or narration/comprehension in editing. Simultaneously, the sound and shots assembled into a film provide multimodal meaning. In film editing, all these aspects, plus other aspects, such as acting and genre, must be adhered to, whether you consider editing an art or a craft, or both.
Perception and Attention in Film Editing

Regarding perception, it is the links with attention that are of importance here. Attention is the function of priority amongst sensory inputs (Lamme 2003). Different kinds of thinking give rise to an individual’s intentions. These intentions have obvious impact on attention (Fry 1981, Hamlyn 1996, Crary 1999). Attention, in turn, directs perception, which picks up stimuli properties (a few, or a chunk of details) and delivers this information to be used for some sort of intellectual or bodily response (see Figure 12). This is the top-down function of attention.

The bottom-up function of attention is that perception – one or several of the senses – reacts to the environment and selects some information that is possibly urgent, and pushes this information into attention, for further response (intellectual, or bodily) by the individual (Pylyshyn 2003, Goldstone et al. 2015), (see Figure 12). The bottom-up and the top-down function of attention operate separately from each other (Pinto et al. 2013), and when details stand out (as transients) the bottom-up function is stronger, and grabs attention. Attention operates in a sequential manner, processing one stimulus (or phenomenon) at a time, and is conditioned by the configuration of the stimuli properties (Egeth & Yantis 1997), (see Figure 15). Coupled to film editing, the shape of edits matters to the human attention capacity, to balance the top-down and bottom-up functions, attention to the story, or provocation by audiovisual transients (Geiger & Reeves 1993, Batty et al. 2015, Cohen et al. 2015). Continuity editing is attention grabbing, engaging viewers so that their focus is caught and withheld during the run of the film, which was what Münsterberg (2002 [1916]) claimed. In this context, the multimodal effect of sound with image as strengthening attention must be recognized (Shimamura et al. 2015).

Figure 12. Model of attention, and how it relates to intention and perception.
Visual perception, in the Gibsonian understanding, has implications for design (see Figure 13). Gibson’s (2015[1979]) ecological approach to visual perception takes the embodied being as a whole into account. By changing position, the individual can alter points of observation in regard to the environment and objects in it. From each position, the individual has a direct perception of what is seen. The object properties that do not change with point of observation are invariants. An object’s invariants constitute its affordances.

An affordance, in turn, can be a capacity-possible-to-employ by the individual perceiving it, as I understand it. This way, the individual interplays with the environment and perceives its affordances (Turvey et al. 1981). Affordances have also been claimed to be dispositions, dependent on the abilities and deeds of the individual (Scarantino 2002), or even as instances of culturally-derived functional meaning (Heft 1989). For example, whether or not an individual is familiar with what a TV-screen is, determines if s/he perceives its affordance to show a lit picture, and how it functions in terms of being switched on or off. This is how I understand Norman’s (1998) idea that a user should directly perceive designed items (and their functions).

Regarding film editing, Gibson states that:

The art of film-editing should be guided by knowledge of how events and the progress of events are naturally perceived. The composing of a film is not analogous to the composing of a painting. The sequential nesting [Gibson’s emphasis] of subordinate events into superordinate events is crucial. The transitions should be psychologically meaningful, and the sequential order of happenings should be intelligible. [...] The various kinds of filmic transitions – zoom, dolly, pan, cut, fade, wipe, dissolve, and split-screen shots – could usefully be evaluated in the light of ecological optics instead of the snapshot optics that is currently accepted.


This notion of how events are perceived in film, as major and minor events that a film viewer can understand, is an approach Magliano and Zacks (2011), as well as Cutting et al. (2012), adopted when they study event perception in film editing, addressed below, and under Perception Research. The general relevance of Gibson’s theory of ecological perception to film viewing lies in the directness of the functioning of perception, which could possibly be disturbed by film editing that does not support the fluidity of event evolvement during the run of a film.
Figure 13. Model of Gibsonian “direct perception”, ecological interplay, and relations between affordances and design (my interpretation). Bodily senses perceive (the objects of) the environment, and what does not change with different points of observation are invariants. Invariants constitute an object’s affordances, which can be understood as either dispositions, capacities-possible-to-employ, or functional meaning approaches that can be employed when designing artefacts so that their affordances, and any action required, are immediately perceivable for its users.
Visual perceptual responses to visual phenomena are possible to capture and measure, by means of eye tracking. What is captured are moments of relative stillness of the eye, called fixations, and eye movements between fixations, called saccades. Fixations are measured for their position (e.g. on a screen) and time duration, while saccades are measured for distance (between fixations), and occasionally for time and speed. Long saccades indicate a change in point-of-interest regarding the visual phenomenon, or stimulus. Short saccades occur within a limited point-of-interest, in combination with fixations (Buswell 1935, McCarley & Kramer 2007). Together, fixations and saccades across a visual stimuli form a gaze pattern (see Figure 14), a “local” pattern for a point-of-interest, or a “global” pattern for the whole stimuli (see Figure 15). Given tasks strongly influence viewers’ gaze patterns (Yarbus 1967).

![Figure 14. Gaze Pattern consisting of fixations, indicated by rings where ring size is proportionate to fixation time, and saccades, represented by lines. Long saccades indicate a change in point-of-interest, whereas a short saccade is a movement within a point-of-interest (Buswell 1935).](image)

The gaze pattern associated with a given task reveals what has been attended to, as well as in what order (see Figure 15), (Yarbus 1967).

![Figure 15. Tasks influence viewers’ viewings of visual stimuli, which can be captured in gaze patterns. Gaze patterns (either local or global) tell something about to where the visual attention has been directed.](image)
The film creators’ efforts to guide viewers’ gazes are confirmed by Marchant et al. (2009), as well as the correlating conformation (design) of filmic expressions to viewers’ gazes, found by Cutting et al. (2012), are found to match viewers’ appreciation of filmic events (Cutting et al. 2012). A similar adjustment of film edits to unfolding events was found by Magliano and Zacks (2011), where continuity edits were related to event starts and endings (see Figure 16). Continuity or discontinuity are editing expressions (cf. Orpen 2003) that are formed by means of linkage or collision, and support experiences, such as rhythm (Pearlman 2009).

![Diagram](image.png)

**Figure 16**. A film’s design directs viewers’ gazes (visual perception), which match their detection of filmic events. Event segmentation is adhered to in continuity editing, whereas counteracting event segmentation supports the editing expression of discontinuity.

Continuity edits and discontinuity edits have been found to influence perception, attention, and cognitive processing in different ways (d’Ydewalle & Vanderbeeken 1990, Geiger & Reeves 1993, Hochberg & Brooks 1996, d’Ydewalle et al. 1998, Smith 2005, Smith & Henderson 2008, Wang et al. 2012). Visual queries and answers enhance continuity (Hochberg & Brooks 1978), as do established continuity rules (d’Ydewalle & Vanderbeeken 1990), and similarity in content (Geiger & Reeves 1993). When continuity is disrupted and replaced by discontinuity, perceptual reactions increase (d’Ydewalle & Vanderbeeken 1990), and attention is drawn to the edit, increasing the awareness of editing (Smith & Henderson 2008). Consequentially, cognitive load increases (d’Ydewalle & Vanderbeeken 1990, Geiger & Reeves 1993), leading to decreased understanding and poor memory of what is viewed (Smith 2005),
(see Figure 17). Continuity edits instead support these cognitive functions, and continuous attention to the film story (Smith 2005).

**Figure 17.** Influence model of continuity and discontinuity edits on perceptual reactions, attention, awareness, cognitive load, memory and understanding. Arrows indicate direction of influence, where + stands for an increase at that end of the arrow, and – represents a corresponding decrease. Discontinuity increases attention and awareness of the film as a construct and its edits, and decreases memory and understanding. Continuity has the contrary effect. (The model represents my interpretation of d’Ydewalle & Vanderbeeken 1990, Geiger & Reeves 1993, Hochberg & Brooks 1996, d’Ydewalle et al. 1998, Smith 2005, Smith & Henderson 2008, Wang et al. 2012.)

Explicitly, how perceived continuity is supported by continuity editing is researched by Smith (2005). He suggested that it is the removal of perceptual disturbances, called visual transients (cf. Germeyns & d’Ydewalle 2007), that supports time continuity expectations, as well as space continuity expectations. When viewers’ expectations on spatiotemporal continuity are high, they also exceedingly expect objects to continue to exist before them, and as a result they are more disposed to perceive continuity regarding the film as a whole (Smith 2005), (see Figure 18).
Figure 18. Influence model of how removal of perceptual disturbances in continuity editing supports viewers’ perceived continuity (interpretation of Smith 2005). Arrows indicate direction of influence, where + stands for an increase at that end of the arrow, and – represents a corresponding decrease.

Berliner and Cohen’s (2011) model of how the mental construction of the illusion of continuity of the real world functions is supposed to have a parallel functioning for film viewing (see Figure 19). Fragmented and disjointed visual inputs are used, and information gaps regarding the world are disregarded, when a mental model of the world as a “whole” is created. In a similar fashion, a filmic world is created from the fragmented audiovisual inputs the film provides, supported by continuity editing.

Figure 19. An interpretation of Berliner and Cohen’s (2011) model of how an illusion of continuity of the real world is created as a mental model, from fragmented visual inputs, and gaps. Perceiving a film world has its parallel in a similar mental treatment of audiovisual inputs.
Model Application and Interplay

The theories I use, which add to the interpretation of my data, only clarify it partially, and from specific perspectives. Therefore, in this section, I will discuss how these partial clarifications relate to and interplay with each other, how they interplay with other theoretical concepts mentioned, and how they relate to my research questions. The application of the theories in interplay during the analysis will be described explicitly, as well as purpose, and expected outcome.

When applying the above described theoretical models, I will turn to my first research question: What does perceptual precision attain in the design of continuity edits and for viewers’ film decoding? This question assumes that film editing could be addressed as design. However, I will support that assumption by letting some of the presented theories interplay into a Design Profession Interplay Model (Figure 20): Murch (2006) ascribes film editing with the property of being a silent and unrecognized profession. Smith and Whitfield (2005) make a similar claim about many design professions, and suggest that each profession could become easier to recognize from a public perspective, if identified and described by a topography of characteristic work achieved by practitioners of that profession. The film editing literature provides candidates for such topography: (Lindgren 2011[1948], Reisz & Millar 1968, Murch 2001, Pearlman 2009, Billing 2017). According to this literature, film editing consists of selecting shots from a stock of footage, ordering the chosen shots into a sequence, arranging sound in regard to that sequence, and trimming edits for the desired kinds of transitions, all meanwhile achieving an audio-

![Diagram](image)

**Figure 20.** A Design Profession Interplay Model: design professions can be understood from different perspectives: from within the community of those holding the profession; from an academic interest in professions; and from the holders of public agencies. The public perspective relies on the academic perspective, which in turn relies on the “within” perspective. Typical actions are of central interest in all three perspectives.
visual flow with rhythmic properties that the film editor masters. Such characteristic work could constitute essential parts of the film editor’s profession topography.

In Paper F, I address a film editor’s work by these categorized activities, for further issues regarding the use of perception in documentary film editing. This academic attempt to understand film editing is engaged with the formation and application of audiovisual knowledge and skills, as exercised through a profession. In trying to distinguish this profession, Orpen (2003) places film editing in between art and craft. It is a craft in accordance with the use of routines and conventions in accomplishing an audiovisual object, and an art as it strives for expressiveness. Moreover, Niedderer and Townsend (2014) suggest that crafts should be studied with a design approach where crafting activities are at the core of their interest. It is the particular manners in which a craftsman becomes involved in these activities that clarify the formation of skills and knowledge of the craft.

For further understanding of design professions, as designers understand their profession from within, Adams et al. (2011) suggest that a designer’s actions must also be acknowledged, along with the self-understanding of being a designer, as well as specific knowledge formulated by the designer. Many of the writings on film editing are authored by individuals with a background in film editing, and incorporate such understanding and knowledge. By letting this formulated understanding of film editing, and examined documentary film editing work (Papers E and F) be subjected to academic analysis, it can be transformed into a profession topography, from which film editing may be recognized and regarded as less silent as an audiovisual design profession.

In support of this reasoning, Paper A, along with Westera (1999), considers film and TV production as a whole to be engaged in audiovisual design. Further, similarities are shared with other design contexts, where aesthetic considerations for user attraction and engagement also manifest themselves during the material treatment of designed objects (Paper B, Lewalski 1988, Bloch 1995, MacDonald 2001, Jordan 2002, Sato 2009). The next aspect of that same research question I will address regards the very design of film edits: the material handling of the shots and sounds that meet at an edit point. Like design in other contexts, audiovisual design can be approached as process, including sub-processes, during which materials are treated, design problems addressed, and design decisions taken, for the making of an artefact (cf. Westera 1995, Eckert & Clarkson 2005, Poggenpohl & Sato 2009). Such material handling is discussed by Brinck as aesthetic problem-solving (1999 & 2007). Poggenpohl (2009) suggests that problem-solving and meaning-making integrate during a design process, with the emphasis on problem-solving at the earlier stages, and meaning-making at the later stages of the process.
In the film editing process (see the Fead-model, Figure 21), the desired outcome is an artefact – a film (sequence) – that is meaningful to its audience, and makes some sort of sense from some respect. At the start of the process there is a stock of (digital) material that is to be compiled into this artefact. The design problem is how this can be accomplished: how can the story (or other content) be better composed for intended meaning and comprehension? What shots and sounds support that composition? How can these be better ordered? How can they be better matched, to merge or contrast? How can the editor make sounds and shots merge, as film continuity, or contrast, as discontinuity? Should linkage or collision be used? These questions stack up for a problem-solving sequence, and are mostly aesthetic, where design decisions are made. This is how film editing includes framing and solving design problems, through design decisions, while making meaning.

The material itself is conditioned by being of a documentary kind (Rabiger 2009, Reisz & Millar 1968). The ability of the material to express things, as well as the extent to which it can be processed, with preserved credibility, is limited (Manovich 2001 & 2012, Eriksson 2013, Eriksson P.E. & Eriksson Y. 2015). The compilation of an audiovisual object includes the problem of matching shots and sounds at the edit point, which depends on the material,

**Figure 21.** The model of film editing as an audiovisual design process, Fead, consists of framing the design problem, solving it through design decisions, which along with the work renders audiovisual meaning (potential) for film viewers when they meet the artifact: the film. The film will consist of audiovisual (digital) material that is compiled into an audiovisual object by many sound and image components. This compilation must regard the material’s expressability, its processing, its audiovisual matching, and its semiotics (potential). The editing aspects of the film as an artefact will have properties that appear as audiovisual merge, linkage and collision, continuity or discontinuity.
and how difficult it will be to achieve continuity, or discontinuity, and relate to that how the emotional response to edits will be: smooth or harsh (Reisz & Millar 1968, Westera 1995, Orpen 2003). Some semiotic properties that sound and image features carry could jointly point to unwanted meaning so that they multimodally communicate what was not intended. The coherency of the audiovisual object depends on how well audio and visual merge, and support a unified affordance of meaning (Westera 1995, Kress & Van Leeuwen 2001, Ponte & Niemeyer 2013). These aspects of the design of each edit must be regarded collectively before arriving at the perceptual precision aspect of the research question. They are expected to relate to the formation of continuity or discontinuity edits in the artefact. When images graphically link with each other, rather than collide, film continuity or discontinuity is achieved (Reisz & Millar 1968, Orpen 2003, Smith 2005, Pearlman 2009). Continuity is known to be crucial to narration and comprehension, as reception instances (Smith & Henderson 2008, Berliner & Cohen 2011).

Support for this view is found in research concerning design processes, design issues and problems, and design work, and its various aspects, components, activities, and sub-processes (Eckert & Clarkson 2005, Poggenpohl & Sato 2009). Some design research also relates design processes to design principles in use in specific contexts of particular rationale (Horváth 2004, Pahl et al. 2007, Blair-Early & Zender 2008, Poggenpohl 2009).

I will now turn to how continuity and discontinuity relate to perceptual precision and viewers’ film decoding (presented as the FIPA-model, Figure 22). Perceptual precision means (a designer’s) close consideration of details that disturb perception and threaten attention. When viewers perceive a continuous filmic world, as an illusion, this is their mental accomplishment from fragmented audiovisual inputs (Berliner & Cohen 2011, cf. Figure 19). Perceived continuity is based on the expectations of the constant existence of the filmic world and its inhabitants, which, in turn, is supported by spatiotemporal continuity, i.e. film continuity (Smith 2005, cf. Figures 16 and 18). Perceptual disturbances, i.e. audiovisual transients, may jeopardize the spatiotemporal continuity. Continuity carries few and modest transients, and hides them well, while discontinuity comes with many or palpable transients (Smith 2005). Transients are known to provoke perceptual reactions (d’Ydewalle & Vanderbeeken 1990), affecting attention (Smith & Henderson 2008). These effects are also coupled to increased cognitive load, which lessen understanding, as well as weaken memory (Smith 2005). The film editor’s perceptual precision is expected to support the accomplishment of continuity edits, if so desired (cf. Figure 17).

The second research question of this thesis: *In what ways could (audio–) visual intentions be dependent on perceptual precision in the design of continuity edits?* is contingent on the answers to the first research question. When we know the influence of perceptual precision, we can relate it to design intentions.
The film editor’s intentions are enfolded both in the purpose of the documentary film sequence — to be a comprehensive narrative (cf. Berliner & Cohen 2011) — and in her thinking on specific sounds, shots, and edits. When she intends continuity in support of the narrative, or specific audiovisual meaning, these intentions can be reinforced by means of her attendance to (audio-)visual perceptual phenomena (Smith 2005). This attendance to (audio-)visual perceptual phenomena during editing can be related to perceptual responses of the viewers, in terms of gaze patterns (Hochberg & Brooks 1978 & 1996, d’Ydewalle & Vanderbeeken 1990, Geiger & Reeves 1993, d’Ydewalle et al. 1998, Smith 2005, Smith & Henderson 2008, Wang et al. 2012, cf. Figure 14). The viewer’s task of following the narrative through the course of the film is dependent on continuous comprehension, which in turn is dependent on continuous unconstrained and direct perception of what is central and meaningful to the story (cf. Figures 13 and 15). Whether the film editor’s intentions are dependent on attendance to perceptual phenomena, in terms of its influence on viewers’ perception, can thus be discussed.

Figure 22. The FIPA-model of film (as an artefact), (the film editor’s) intention, (viewer) perception and attention, connected to the theories of influence of tasks, interpretation of gaze patterns, continuity editing and film continuity, Gibsonian direct perception, perceived continuity, the illusion of continuity in film experience, and continuity support for understanding and memory (cf. Figures 12-19). The decisive connection point is how perception meets stimuli properties.
Methods

The systematic approach that scientific research is expected to obey requires a sincere account of procedures and tools utilized when acquiring and analyzing data. Such an account provides the justification of the results, an assessment of benefits and drawbacks of the techniques and processes used, and offers other researchers the possibility to repeat or improve the means of inquiry. This account of the ways research has been conducted is the methods.

My methods chapter is divided into two main parts: the principal research disposition is presented in the chapter Research Design, where the methods used in the different studies are presented, followed by the data categories generated, and then operationalization of concepts. Thereafter, Methodology and Reflections, as a separate chapter, carry a method criticism where the different methods that have generated the data for the thesis, are discussed more elaborately.

First, however, I will address some issues regarding the scientific tenets of this thesis:

Design calls for a change for the better (Simon 1969). It wants to be meaningful to users (Krippendorff 1995). Design research also calls for influence. It wants to be applicable to design work in the contexts where it prevails (Blessing & Chakrabarti 2009). This thesis calls for improvements to occur in design contexts where (audio-)visual perception is applied by designers to shape (audio-)visual meaning. It wants to improve (audio-) visual design processes where perceptual precision is a factor that matters to designers, such as film editing students and inexperienced film editors, as well as other designers that use (audio-)visual material.

In my view, academic research must start with a meticulous framing of a research problem as a lack of knowledge of some kind. This lack of knowledge is what is addressed through a meticulously formulated research question, where the answer to that question will, in the best case, amend the formulated research problem, or, at least, do so in part. The crucial thing is where to find the answer. But the answer does not, generally, hide in a place waiting to be uncovered. Rather, the answer to the research question is extracted through the analysis of a dataset. The researcher’s duty is to understand what categories of data, and how much, is needed for this extraction. What can be extracted from what kind of data? Thus, data might be needed that expresses ideas, thoughts, or understanding: e.g. words, moving images, or sound – that is, data of a qualitative nature. Likewise, data that expresses measurements of
extent or amount (figures) might be necessary – that is data of a quantitative nature. Hence, addressing a particular research question might require a dataset consisting of several kinds if data, of both a qualitative and quantitative nature. The researcher’s means for rendering the dataset necessary is the method.

Understanding a dependency, which my first research question addresses, must appear as a proposition, since an academic understanding of something entails being able to formulate its meaning. Thus, the answer to that question has to be of a qualitative nature – arrived at by means of reasoning. Likewise, the second research question, addressing the influence of a phenomenon in a certain context, must be formulated as a proposition. The answers to both research questions can be supported by an assessment on some sort of scale – that is, by data of a quantitative nature. Thus, the answers to the research questions can be based on a combination of qualitative and quantitative data.

The data underpinning these answers can exist as wordings (qualitative), or as figures (quantitative) (cf. Åberg 2001). An influence indicates relations and a context in which it resides. Gathering data about the influence some phenomenon has, therefore, requires capturing the context and relations between the phenomenon and the context, during its emergence to that point. Some aspects of a phenomenon might be measurable, or quantifiable, and other aspects might totally escape any tool of measurement, but might appear in other intelligible forms that can be captured and subjected to judgement regarding its qualities. To estimate the influence of a phenomenon in its context implies adopting a comparison of the phenomenon to other phenomena in that context, for assessing their influence on each other, or on the context.

Considering the scientific rigour of the data, this should have truth value, consistency, neutrality, and applicability (Hamberg et al. 1994). Quantitative data gains its truth value from its validity, whereas qualitative data must be credible. Quantitative data gets its consistency from being reliable, while qualitative data has to be dependable. Quantitative data is neutral when it is captured in an objective manner, whereas qualitative data must be possible to confirm. Quantitative data is applicable if it is useful for generalization, forasmuch as qualitative data is required to be transferable (Hamberg et al. 1994 p.178).

Suggestions for how phenomena can be considered in coming situations are speculative. However, if the suggestions emerge from, and are well founded in, understanding, and in the evidence they are based on, the suggestions might still be the most well-suited improvement from the current situation (Blessing & Chakrabarti 2009), as a form of transfer. The results of this thesis are thus usable in understanding film editing, but may be transferable to other (audio-)visual design contexts.
Research Design

The understanding of research results, and whether they are transferrable to other contexts, depend on our accepting the design and set-up of the research methods. In this chapter, I will describe how the studies on which this thesis is based upon relate to the thesis and to each other. I will also describe the different aspects the studies address, as well as why these aspects are important as a foundation to the thesis. The methods used will be outlined and related to the data generated, and how this data can be used to answer the research questions will be described. Prior to that, I will now summarize the methods used as my research outline:

The research field of film and TV production is clarified, in part, by a literature study scrutinizing experienced film and TV professionals’ writings on industrial film and TV production processes, and, in part, by a study where managers of film and TV production were interviewed regarding the shift of technology and following production cruxes. These interviews were analyzed for key statements regarding process management and workflow (Swenbergen 2012, Papers B and C). In the literature study, the analysis regarded statements on film and TV production processes, issues and problems, as well as particular work tasks during the creation and processing of sound and images, and whether these statements indicated any characteristics that could be usefully approached from a design research perspective (Paper A).

Thereafter, the field of film editing is explored by setting up a study (preluded by a pilot-study) in which a method is developed that can capture a film editor’s intentions while editing a documentary film sequence. This is done through video observations by a web-camera and by running computer screen recordings of the editing software (Papers D and E). I also conducted an elicitation with the film editor, where she was asked to present her thinking on the editing of the film she had just completed. This video elicitation was likewise recorded via web-cam and running screen recordings (Paper E). Subsequently, all video recordings were analyzed for kinds of editing work executed by the film editor, as appointed by editing literature. The film editor’s intentions with sound and image were captured. Time consumption and re-iterations of edit points were counted. Further, involvement of perceptual aspects of the editing work was categorized, in accordance with the tradition of perception psychology.

Next, the understanding from the intervention into the field is examined against viewer tests. Film viewers were eye-tracked while watching two instances of the documentary film sequence that the editor had made: the original one by the editor, and one version that I had manipulated slightly. In the analysis, the editor’s intentions for viewers’ gazes were compared to the viewers’ actual eye movements, to see whether appointed Areas of interest (AoIs) were hit or not, but also to see whether the amount of eye movements would differ or not, as well as whether pupil sizes would be differently affected, or not (Paper F).
Hence, the thesis uses a three-stage research design where, first, film and TV production is examined for its suitability for a design research perspective, second, a documentary film editing process is examined for intentions and use of perception, and, third, viewer responses to film edits made are analyzed.

Three Studies – Methods Outline and Interrelation

This thesis is based on three studies and a pilot-study. Their respective method will be presented below, and their interrelation explicated.

The first study

First, I carried out a literature study of film and TV craft practitioners’ writings on their experience with film and TV production, that addresses film and TV production as to whether it could be regarded as design (Paper A). The literature search was aimed at books that specifically cover over-all production processes in the film and TV industry (a topic too wide to be covered in papers), written by experienced and renowned practitioners from the film and TV industry, who were well familiar with the processes, issues, and problems that repeatedly occur within such production. The bulk of such literature is in English, so therefore the search was done using WorldCat, Summon, Library of Congress, and the EBSCO Film & TV Literature Index, with the language of origin limited to English. Search terms used were: “film production”, “TV production”, “television production”, “moving image production”, and “motion picture production”. The authors’ statements on film and TV production were then analyzed in regard to design theory, in order to examine whether there are characteristics of such production that might benefit from being studied from a design perspective. Film and TV production is herein delimited to the material handling, creation, or processing, of sound and image.

The analysis focuses on three topics:

1. how to understand what can be framed as design, or not, during the material handling of sounds and images in film and TV production;

2. what can be gained by addressing film and TV production processes from a design perspective; and

3. what support design research can provide in understanding film and TV production issues and problems.

The study provides arguments for considering core parts of film and TV production as audiovisual design, which constitutes a design research perspective. From that, processes, issues and problems within film and TV production can be compared to processes, issues and problems in other kinds of design. Perceptual aspects of film and TV production can be compared to perceptual aspects of other (audio-)visual design, as well.
Furthermore, the analysis motivates understanding of the film and TV production processes examined in the first study, presented in my Licentiate thesis (Swenberg 2012) and Papers B and C as audiovisual design processes. Documentary film editing, scrutinized in the third study, is a specific case of such audiovisual design. The research questions explore this context of film editing as audiovisual design, and the answers will relate to it. Overall, I want to look at film and TV production from a design research horizon in this thesis, and I justify that use by means of this study.

The second study
Second, employing a design perspective on film and TV production, I conducted a study much elaborated on in my Licentiate thesis (Swenberg 2012), and in Papers B and C, where the change of recording technology and recording medium in the High End segment of the Swedish film and TV industry was addressed in regard to the design processes following recording. New digital materials had to be handled in new ways, and our analysis of key statements identified factors that linked technological, technical, and procedural issues with management and workflow during creative processing of sound and images. These interviews with managers at a variety of firms with different specializations showed that material factors strangle creativity for audiovisual designers in new ways, due to the new materials and new design processes and workflows.

One of the material factors discussed is that digital audio and video material can be configured in many ways. The resulting bottle-necks and compatibility issues called for new management efforts, which my colleagues and I addressed as a design research effort to improve an existing film and TV industrial situation.

By inventing a concept for an on-line workflow check-list, called MI-Flow (www.mi-flow.se), we could assist film/TV production managers in overviewing file format and codec conversions in a moving image workflow (Paper C). The MI-Flow check-list was tested on members of the film and TV industry to evaluate its principal function against the identified needs for workflow management within the industry. This was done, first, at a group interview setting, and then, for individual testing in an industrial setting with following qualitative questionnaire responses.

These sets of statements from group interview and questionnaires were analyzed according to the same scheme of concepts as the earlier interviews. The confirmation of the principals of the on-line workflow check-list underlines the importance of well-functioning workflows for maintaining creative space during film and TV production. Departing from this position, it is of interest to this thesis to address more in depth factors that enable audiovisual designers, situated in chain design processes, to handle sound and image material more fluidly, if possible. Here, the influence of perceptual precision is expected to come into play.
The third study
In order to address perceptual aspects of film editing I arranged a pilot-study (Paper D) that preluded my third study (Papers E, F, G). The objective for these studies was to address how a film editor’s intentions about the edits made during film editing were perceived by viewers, and whether intended viewings could be established. The full study also included whether the film editor had specific perceptual intentions or exhibited concern regarding perceptual phenomena during editing. Film editing is, from the perspective of this thesis, as well as in regard to the other studies, an instance of audiovisual design. Moreover, since audiovisual design uses multimodal capacities, the research methods must do so as well (cf. Kress & Van Leeuwen 2001, McCarley & Kramer 2007). This study of film editing provides insights into how (audio-)visual perceptual precision is achieved when treating and processing digital audiovisual material, and how that precision matters to film viewers (Paper G).

Video Observations and Screen-recordings
In the pilot-study, the capturing of the film editing process is developed. The process is recorded both with a video camera monitoring the editor from behind, and with a running screen-recording software called iShowU. This software can record all that is monitored on the editing computer screen, including the image from a web-camera facing the editor at work, capturing the upper body. It also records sounds from the computer, including the web-camera’s sound, as well as computer keystrokes. However, iShowU caused the computer used to lag, so for the completion of the pilot-study only the video recordings from the camera behind were used and presented in Paper D.

For the full study a more powerful computer was used, which could easily run iShowU along with editing software, web-camera and other programs, https://vimeo.com/214633391/9a611cfbc1. The recordings of an editor at work are necessary in order to make a post-process analysis of the executed work. A record of keystrokes can reveal what exact functions of the software were used, as well as in what order. The video recordings of the editor reveal the body language and utterances that show when the editor stops to think, or expresses her thoughts or feelings. The screen-recordings show how the editing software is used, how the film material is treated, as well as when iterations and re-iterations are carried out. This data provides evidence about the design process, and about issues and problems encountered by the editor, for the later analysis.

In addition, as researcher, I accompanied the editor at work in the edit suite during the full study, as an observation. During that observation, I could intervene whenever I needed an explanation for the editor’s actions or thoughts. With my professional experience as a former TV editor, I could unconstrainedly follow the process and grasp when there were difficulties. On the final day

64
of editing, we performed an elicitation of the completed film sequence, where
the editor on the first run through the sequence commented freely on her
intentions with the edits and the edit points. Thereafter, we watched the se-
queness again, and I addressed the editor with specific questions that inquired
about different design aspects, but, specifically, about perceptual aspects. This
elicitation session was also video recorded with the web-camera and the run-

Video Analyses
The ensuing analysis of the video recordings considered the kinds of editing
work the film editor pursued, based on their categorization in the editing lit-
erature: assessment of material, work on edits, and configuration of story,
https://vimeo.com/214704609/4f9d6f0cc6 (Lindgren 2011[1948], Reisz and
intentions with the film sequence, its parts, and specific sounds, images and
edits were also extracted from the video recordings of observations and the
elicitation (see Figures 23 – 24).

Figures 23 – 24. The arrow in the left figure shows where the film editor intended
the viewer’s gaze to be in the frame before the 3rd edit (at the left window). The
arrow in the right figure shows where the gaze was expected to land (on a bird’s
nest on the fire-place mantle), the frame after the edit.

Time consumption and re-iterations were also totaled for work on edit points.
Lastly, perceptual phenomena at stake during the editing (as well as commented
upon in elicitation) were identified and accounted for, as due to perception
psychology (Hochberg & Brooks 1978, Smith 2005, Wang et al. 2012). This
account was made by means of perception codes (see Appendix I, Table I),
analyzing the editing process, https://vimeo.com/214726742/2a13e3f7d4, as
well as the resulting film sequence, https://vimeo.com/214681838/ed5f9cfd42.
The fully developed method for this study, and its methodology, VOSMET, is
presented in Paper E (see Figure 25).
Figure 25. Methods progression of the research project *Klippares visuella intention och tittares visuella perception*, presented as the VOSMET methodology (Paper E). Numbers [1-13] indicate the order of occurrence of the different aspects of the project. An existing stock of documentary film material [1] is handled through, and conceptualized by, the editing software [2], which is the film editor’s tool. Tool usage [3] and behavior are captured [4+5], and recorded along with the events on the editing computer screen [7], where the film material is audio-visual. The recordings are analyzed for work events [8] and considerations of perceptual phenomena [10], which are also regarded in direct relation to the full film sequence, after completion. Versions of the film sequence are thereafter tested on viewers, using an eye-tracker [12]. Each analysis delivers some kind of graphical representation of results [6, 9, 11, 13]. (Figure and caption as presented in Paper E).

**Eye-Tracking Film Viewers**

How these editorial intentions were actually perceived by viewers was addressed in a second step of the pilot-study, as well as in the full study (see Figure 25), (Papers D and E). To be able to falsify the editor’s intentions, Condition 1, as they materialized in her documentary film sequence, [https://vimeo.com/210944884/be8001e786](https://vimeo.com/210944884/be8001e786), I made a manipulated version of the sequence, Condition 2, [https://vimeo.com/210953506/9ed3cb3b6e](https://vimeo.com/210953506/9ed3cb3b6e), in which I altered all edits to counteract the perceptual intentions of the editor (see Table 1). Both versions were screened for viewers, using a within-subjects design, in order to avoid individual visual behavior impacting on the results (cf. Holmqvist et al. 2011).
Table 1. Edits, types of continuity, and number of frames altered (incoming, outgoing).

<table>
<thead>
<tr>
<th>Edit #</th>
<th>Types of continuity</th>
<th># altered incoming frames</th>
<th># altered outgoing frames</th>
<th>Sum of altered frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sc A Sp T G</td>
<td>-2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Sc Sp T G E</td>
<td>3</td>
<td>-3</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>Sc Sp T G</td>
<td>3</td>
<td>-3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>Sc Sp T</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>Sc Sp T G</td>
<td>-2</td>
<td>-2</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>Sc Sp T</td>
<td>-2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>Sc Sp T G</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Sc Sp T</td>
<td>-3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>Sc Sp T</td>
<td>-3</td>
<td>-3</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>Sc Sp T E</td>
<td>-3</td>
<td>-3</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>Sc Sp T E</td>
<td>2</td>
<td>-2</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-2</td>
<td>-2</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>Sc Sp T</td>
<td>-3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>15</td>
<td>Sc Sp T</td>
<td>2</td>
<td>-3</td>
<td>5</td>
</tr>
<tr>
<td>16</td>
<td>Sc Sp T G</td>
<td>-2</td>
<td>-3</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>18</td>
<td>-</td>
<td>3</td>
<td>-3</td>
<td>6</td>
</tr>
<tr>
<td>19</td>
<td>Sc Sp T</td>
<td>3</td>
<td>-3</td>
<td>6</td>
</tr>
<tr>
<td>20</td>
<td>Sc Sp T</td>
<td>3</td>
<td>-3</td>
<td>6</td>
</tr>
</tbody>
</table>

Edits 12, 13, 17, and 18 are distinguished as discontinuity edits.

In the pilot-study, a cross-range set of participant viewers was engaged. This was gender-balanced, with ages ranging from 24 to 55, and with educational and film-viewing backgrounds also spread (see Paper D for details). The viewers were seated in front of a stationary eye-tracker (SMI RED60), where they watched a 37-second drama sequence, consisting of 10 shots and 9 cuts, which were considered enough to evaluate the study set-up. The viewers’ eye movements were analyzed for corrections of the gaze after edit points. The results from the pilot-study confirmed the usefulness of the method. It was possible to detect differences in viewings between the versions of the film sequence. The method and its procedure were thus further developed for the full study.
For the full study, 33 participant students were recruited, 30 participants had Swedish as their mother tongue, 17 were females, with an average age of 26 years. Most participants were bachelor’s students, with the exception of 3 master’s students. 32 participants claimed to be inexperienced in film creation, whilst 2 participants attended courses that included moving image production, which 2 participants had attended previously. A third of the participants were students of graphic design, 7 were sound and music reproduction students, while the remaining participants represented a variety of disciplines (including sports management, systems engineering, and human resource management).

In this study, the drama sequence was replaced by a documentary sequence, Condition 1, https://vimeo.com/210944884/be8001e786, due to perceptual considerations (see Audiovisual material, and Paper E). I also manipulated the documentary film sequence, in order to counter the editor’s perceptual intentions, Condition 2, https://vimeo.com/210953506/9ed3cb3b6e, before it was screened for the viewers, as they were eye-tracked. Since the frame-to-frame matching of the edit points was the only difference between the versions, and the screening order was randomized, any difference in viewing (for the group) between versions indicates perceptual differences.

The gaze data was analyzed for hits in Areas of Interest (AoI:s), appointed by the editor, as well as for saccades and pupil dilations. Whether or not there were gaze hits in AoI:s reveals if the editor’s intentions were successful regarding where the viewers were supposed to look on the screen/image. The amount and length of saccades reveal whether or not viewers had to shift their attention across the edit, within or outside an AoI. Pupil dilations indicate cognitive load (Beatty 1982, Kramer 1990, Seeber 2013, Chen et al. 2016), which shows to what extent edits were exhausting to perceive for viewers. These three measures of perception will be used to discuss perceptual precision.
Data Categories

My empirical data derives from three studies and a pilot-study presented in separate papers (see also sections Research Basis, Research Design, Three Studies – Methods Outline and Interrelation, and Methodology and Reflections in this thesis). How these studies contribute to answering the research questions of this thesis is outlined in the Results chapter.

In this chapter, I attend to the data produced during the studies underpinning this thesis, in accordance with the different data categories. Data categories are discussed as such: propositions, audiovisual material, actions, and eye data, in regard to their respective scientific rigour (cf. Hamberg et al. 1994), not according to study. Scientific rigour is observed in relation to research questions, methodology, and techniques for data gathering.

Propositions

Propositions and descriptions are qualitative data in the form of statements that express opinions, ideas, estimations, apprehensions and understandings. In my research, I have gathered propositions and statements in the form of writings, answers to questions, spontaneous comments, and discussions. These writings, answers, comments, and discussions derive from previous research and expert knowledge.

When expert writings are analyzed, these writings are considered credible according to the status of the authors, and their publishing with established publishing houses. Misleading or insignificant content is considered improbable. As data in my study, these expert writings are dependable due to their actuality: some works are contemporary publications, while some are more dated. When older and new literature make similar propositions or descriptions, this is considered an indication of something being less mutable over time. The same applies to literature with an older first publication date that has since been re-edited and re-published in later or contemporary editions. Newer literature has the benefit of including later changes in the field. There exist other titles that could have been referenced instead, which would have rendered the same outcome, and thus confirm the data I present. The applicability of my data lies in the selections process, where only film and TV production industry is considered, as well as overarching production processes that span from ideation to a completed audiovisual object. This data is intended to contextualize film editing from a design perspective.

Data in the form of expert answers to questions, either as e-correspondence, in group interviews, or as oral statements recorded on observation videos and elicitation videos, are considered credible in regard to the professional expertise they possess. The following is the way they have been addressed: the expert answers in group interviews and e-correspondence are dependable in re-
lation to the experience of the experts, as well as the different contexts the experts act within. Answers in observation videos and during elicitation are dependable due to coherence with the observed documentary film editing process, with film editing literature, as well as in regard to the film editor’s professional expertise. These expert answers are confirmable from video recordings and archived correspondence. The statements made can be reviewed by addressing individuals of similar expertise with the same questions for comparison. Transferability is indicated by the experts coming from different professions and mixed contexts. The data from group interviews and e-correspondence support the contextualizing of the film editing, whereas the answers from the film editor are used for understanding her thinking about the work she is doing, her intentions, as well as her ways of regarding the audiovisual material.

The film editor’s spontaneous comments during editing are considered credible as immediate responses to the ongoing process. As long as the comments occur spontaneously, there is no reason to question their trustworthiness. Spontaneity and actual statements can be confirmed through the video recordings. What is stated can be compared to other editor’s thinking in the same situation, and thus further confirmed, and evaluated for transfer. This data is used as above.

Discussions occur during group interviews, as well as in conversation with the film editor. Statements during discussions tend to be more polarized, for better or worse. Possible positions become more outspoken, whereas it is less certain that such positions are actual positions, or just rhetoric. Therefore, such statements are regarded as suggestions to be compared to other data for confirmation and transferability. A comparison can reveal established “truths”, just as well as unknown variations in attitudes, and new ideas. The data from discussions with the film editor was used to design the study, while the data from group interviews was regarded for the film editing contextualization.

Audiovisual material

The audiovisual material is regarded as empirical (qualitative) data in three instances: first, it is documentary film footage (with sound) consisting of a stock of diverse shots that have certain material qualities, wherein perceptual phenomena reside; secondly, the film editing process produces an emerging documentary film sequence, which contains a story with structure and content, as well as a number of edits; and thirdly, the completed documentary film sequence is turned into a stimulus for eye tracking, and, as such, it is also altered for comparison at the edit points in a second version.

The material quality of the actual documentary footage cannot be surpassed when regarding the documentary film editing process. It circumscribes the editing and provides conditions for the appliance of editing principles, such as continuity and discontinuity, by means of the perceptual phenomena contained within the material. Usually, documentary films cannot be, and are
not, shot with a predicted intention of which exact shots are supposed to be matched up with other shots in the subsequent editing process. Rather, film cameras (as well as video cameras) are operated primarily to capture the most important event that takes place at that moment, secondarily to capture any other event or feature of interest, in the order of their rank as interesting objects for the intended story (Nichols 2001, Hampe 2007, Rabiger 2014).

Therefore, any consideration of achieving fluid editing and smooth edits is usually neglected at the documentary film shooting stage. The rather scrambled stock of footage such shooting produces offers multiple challenges for the film editor, if the film is supposed to run smoothly with edits disguised, in accordance with film continuity convention (Bordwell, Staiger, & Thompson 1985, Orpen 2003, Billinge 2017). In this study, such editing challenges help in accenting any struggles the film editor faces, when trying to achieve smooth edits. Whatever is the problem in making an edit smooth, will be materially accessible for analysis.

The emerging documentary film sequence provides data on re-iterations, and reasons for the re-iterations: what is added, or removed, and what was unsatisfactory. Depending on what choices the film editor makes, the un-edited film material emerges as the film sequence, and the properties of the audiovisual material are, thus, either, open for analysis, or for possible elicitation later on.

The completed documentary film sequence must be regarded for evaluation of what editing principles were finally applied, and what perceptual phenomena are at stake at each edit. This data, in turn, is the basis for the altered version of the film sequence. The making of an altered version, in turn, is wholly dependent on the access to the original film material, and is much facilitated by access to the same computer, software, editing project and timeline as the original version used, just making a digital copy of it before the altering.

The credibility and dependability of the documentary film footage lies in the fact that it was produced for an actual documentary film. This is possible to confirm by seeing the film, via the film editor, who also was the producer of the film, and through existing writings regarding the completed film (Jonsson Wallin, Mattsson and Sverrisson, 2013).

Considering credibility and dependability for the documentary film sequence produced for this study, the film editor was requested to make a sequence she was content with, which was to be screened in public with her credentials included. This was a measure taken to ensure she would not have to settle for an inferior result. The quality of the sequence is possible to confirm by letting other film editors evaluate it.

The manipulation of a second version of the film sequence was carried out by me, according to a protocol where the number of frames altered for each edit is recorded. The number of frames altered is chosen in regard to undoing the film editor’s intentions with the edits in the original sequence. Altering of the edits can be confirmed, knowing these intentions.
Actions

Actions and deeds are data, captured on video recordings in two contexts: the film editor’s work is recorded as a running screen recording where a webcam monitoring the film editor and myself (observing) is visible. This data can be analyzed both qualitatively, for its work task categories and functions, as well as quantitatively, for occurrences, time consumption, or order.

The documentary film editing process consists of actions and deeds recorded by means of a computer with video editing software. By use of this software the audiovisual material is sorted and handled. These actions are what the editing work consists of, which is part of what is to be analyzed.

There is high credibility in the video recordings of the film editor at work: any action she takes is noted by the camera and recorded from the screen as interaction with the editing software. Any action taken with the audiovisual material is also captured, as are all key-strokes, and body movements. The dependability of these actions is lower, since these recordings concern only how this one film editor acts during editing work. To what extent her actions are transferrable to other film editors must be assessed via similar recordings of them. To what extent the film editor in question here uses relevant actions can be confirmed by the film editing literature.

Eye data

Eye data is quantitative data from eye tracking in a study of film viewing. When a human watches stimuli (e.g. film) on a screen, the eyes are estimated to follow whatever is attended to, that is, monitored on that screen (Duchowski 2007, Holmqvist et al. 2011). Eye data is measurements of in what direction eyes are pointing when looking at a screen — coordinates on the screen surface: x for horizontal position, y for vertical position, as well as what size each pupil has at every moment. Changes in pupil size can be used as an indication of cognitive load (Beatty 1982, Kramer 1990, Chen et al. 2016). For gaze positions, each measured coordinate is given an x and a y value. These coordinates can be related to how a visual stimulus is spread over the screen surface, since we know the size and spread of it.

Eye data was recorded by means of an eye-tracker using the pupil-corneal-reflexion method, which gives an objective measurement of eye positions at a rate of 120 samples per second. Validity is achieved through calibration of the equipment for each participant, which is done by a nine-point calibration, centre and edges, on the screen where the stimulus is to be presented. Each calibration was validated against a maximum-1°-of-visual-angle-deviation criterion. Reliability depends on the participants, and all participants were adults, who are understood to have consistent eye behaviour. Participants were recruited to have a spread of properties in the film editing pilot-study,
and as more homogenous groups in the other studies. Participants were controlled for normal or corrected-to-normal eye sight, and (with the exception of the pilot-study) they were introduced to the experiments by settling them down to obtain a generally relaxed attitude. Pre-experiment knowledge of the purpose of the experiment was avoided, in order not to affect viewings, since viewing is known to be very task-dependent. Instead, guiding tasks were given to obtain normal viewing attitudes. Generalisation also depends on the participant group, so our results are of an indicative kind since most participants were university students.

The recorded eye-position samples were then translated into eye movement behaviour, as fixations, where the eye is relatively still – however, almost never completely still, and saccades, when the eye travels between fixations. Within a fixation, the eye is allowed to move as much as 0,8° of visual angle before it is considered to have relocated. Saccades were also distinguished by two categories, short and long: Short saccades are adjustments of the eye position within a larger point-of-interest in a stimulus, e.g. a mouth in an image of a human face (cf. McCarley & Kramer 2007). Such saccades are comprised of movements on the scale 2°– 8° of the visual angle.¹ Long saccades are movements between points-of-interest in a stimulus, e.g. between a mouth and an eye in an image of a human face. Such points-of-interest may belong to the same analytical Area-of-Interest (AoI) of the stimulus, e.g. the face of a human being, represented in an image. During a long saccade, the eye travels a minimum 8° of visual angle across the stimulus.

There is also a third possible category of eye movements that could be distinguished when regarding film viewing, namely, smooth pursuits, which are the movements the eye makes while following a moving object at a moderate pace across the screen. Yet, analysing for this kind of eye behaviour was not considered to add anything significant to whether or not film viewers hit the film editor’s intended point-of-interest within the moving image frame, after an edit. Thus, smooth pursuits are not considered, per se, in this thesis, but interpreted as fixations and short saccades.

Pupil dilation has become an established measurement of cognitive load, whereas it is known that many factors can have influence on pupil reactions (Beatty 1982, Kramer 1990, Chen et al. 2016). Therefore, pupil dilation can be used to indicate general cognitive load (Kramer 1990), or one must limit possible factors to one, in order to trust the results. A recurring use of pupil dilation as measure of mental workload is when testing mathematical problem-solving (Beatty 1982, Chen & Epps 2014), and configuration memory (Chen, S., Epps, Ruiz, & Chen F. 2011, Chen & Epps 2014, Peysakhovich, Causse, Scanello, & Dehais 2015). In particular, pupil dilation is a stable measurement with young adult participants (Paas, Touvinen, Tabbers, & Van Greven 2003).

¹ For analytical reasons, there is an orthodox gap between the maximum size of eye movements within a fixation, and the minimum size of a short saccade.
Operationalization of Theoretical Concepts

When the gathered data is discussed, some terms and concepts (may) describe or explain aspects of the data. This includes categorizations and measurements. In this section, these terms and concepts will be operationalized as appropriate, with explanations to how they are applied, as well as what is clarified by their appliance.

Emotion and experience are regarded by Niedderer and Townsend (2014) as central aspects of understanding craft activities. Experience is regarded as objective, possible to scrutinize, and when regarding film editing, experience is categorized in terms of what editing principles are applied during editing, per edit, namely, either continuity or discontinuity. These categories are open to qualitative measurement in terms of time-continuity, space-continuity, object-continuity, or none (Smith 2005). The film editor’s semiotic considerations can be objectively established, where I apply the categories of sound and image details. Emotion is intrinsic to the individual, thus not open to measurement, but film edits can be felt as being harsh or smooth, categories I will use as expressing emotion.

Material quality has emerged as an important aspect of moving images. Here, I will consider the processing of material as a category during film editing, where processing time and number of re-workings (iterations) of an edit are the measures. The processing includes perceptual aspects, semiotic aspects, as well as other work task categories (see next page). The material quality is considered to differ between documentary film and fiction film, from the point of aesthetic considerations. The categories of planned or un-planned: story, shots, sounds, and matching are used to distinguish between these aesthetic qualities of the material. A qualitative assessment of shots can distinguish whether or not they are apparently made to match each other in editing. The material may also contain audiovisual details that disturb viewers’ attention, and provoke perceptual reactions, and possibly attention. These details are categorized as transients. Viewer perceptual reactions to film edits are possible to measure via eye tracking, as observed perceptual responses coupled to attention (see perception and attention).

Perception and attention are linked categories, where attention is found to follow the gaze rather closely (Duchowski 2007, McCarley & Kramer 2007). Especially, this is the case when individuals are given a visual task that includes a visual stimulus, as participants are in an eye-tracking lab. Attention is the category of interest to study, whereas the gaze is what we can measure in terms of eye fixations and saccades (relocating movements). For fixations, the time duration and location on a screen are measured. For saccades, the distance (between two fixations) is measured. Over time, the gaze forms a pattern from fixations and saccades (see Figure 14). When this gaze pattern is superimposed upon the visual stimulus in question, gatherings of fixations and short saccades indicate a point-of-interest (not necessarily intellectually interesting) in
the stimulus (Figure 26). Such a point-of-interest is regarded as a locus of attention. Longer saccades bridge between such loci. The resulting gaze pattern can be analyzed in terms of its spread, as an overview (global) pattern covering a large area of the stimuli, or covering a local, minor part of it (see Figure 26). These are qualitative measures.

A series of fixations also represents the time course of attention, where the total dwell time in an appointed area-of-interest of the stimuli (Aol), or the number of revisits to that Aol can serve as measures. Whether or not a viewer hits an Aol is a measure that can be coupled to the purpose of the stimuli, or to what has been appointed as important features of it (see Figure 26). Why the gaze has found a point-of-interest, and located attention there, is considered to have to do with the detail properties of the stimulus. Some detail(s) has grabbed attention, or vice versa (see Figure 26).

**Figure 26.** Gaze pattern (yellow) superimposed on a frame from a stimuli video (the documentary film sequence from the film editing study), representing how the behavior of a film viewer’s gaze may occur during a couple of seconds. Rings represent fixations, where ring size indicates the endurance of the fixation, whereas straight lines represent saccades between fixations. Fixations 2-3, and 4-5-6 are examples of fixation gatherings (points-of-interest). The total spread of the gaze pattern across the frame tells about the viewer’s global attention to the shot, while the local pattern of fixations 1-2-3 tells about his/her attention to a minor part of it (workshop machinery).

Also, the main Area-of-Interest (Aol) is represented by its boundary line (pink). The total viewer attention within the Aol can be measured as dwell time. Numbers of revisits (in this case one, fixation #8) is another Aol attention measure.
Contrary, transient stimuli details can also provoke perceptual reactions that counteract attention, disturbing the gaze in its progress of examining the stimulus. Such disturbances can be measured in terms of increased saccade frequency (number of saccades in a short time interval), as a visual response to the transient (cf. Figure 27), whereas continuous focused attention has a stable visual behavior with mainly fixations as its response. In an analysis, these eye behaviors must be located at the appropriate time interval after the transient or other audiovisual feature occurred.

Furthermore, perceptual disturbances can provoke cognitive load, which, in turn, affects memory and understanding. Considering film perception, edits may provide perceptual disturbances that cause perceptual reactions, which could make attention shift from the film story to other inputs, and thereby affect cognitive load.

![Figure 27. Influence model of audiovisual (AV) transients' effect on cognitive load. An increase in number of AV transients triggers the gaze to move, so the amount of saccades increases. More saccades means more work, whilst they also boost shifts of attention. Higher frequencies of saccades, and attention shifts, both cause higher mental workload (cf. Smith 2005), which is known to affect pupil dilation (Seeber 2013).

A decrease in AV transients increases the stability (length) of fixations, which, in turn directly eases the mental workload. Longer fixations also counteract attention shifts, and thereby, the two lessen the cognitive load.

Cognitive load is also found to coexist with an increase in pupil dilation, which can therefore serve as a measure. A specific option for measuring pupil dilation at film edits is to adhere to the time intervals close to an edit, identified by Magliano and Zacks (2011), as correlated to event segmentation (see}
Figure 27). From tests using mathematical problem-solving tasks, or configuration memory tasks, easier tasks show pupil size increases of approximately 5%, whereas when tasks are difficult, pupil sizes increase by 10%, and more (Beatty 1982, Chen et al. 2011, Chen & Epps 2014, Peysakhovich et al. 2015). Intermediate task difficulty is accompanied by 7-8% larger pupils.

Film editing work tasks are analyzed in the categories of assessment of material, work on edits, and configuration of story, from the editing literature (Lindgren 2011[1948], Reisz & Millar 1968, Murch 2001, Pearlman 2009, Billinge 2017). Configuration of story is further divided into build of story (structure), and choosing story content. Some other work tasks are summed up in a surplus category, namely, other editing work. The measure used in this case is time proportion, which indicates an editor’s topography.

The processing of edits was analyzed specifically for audiovisual perceptual properties at stake during the shaping of the film edits: perceptual phenomena were categorized and counted per edit and compared to time consumption (see Papers E and F). The total time spent on deciding and elaborating each edit point was measured. The number of occasions the editor spent working with an edit was counted, from first deciding on it, and then re-processing its edit point, until content with its appearance. Time consumption and re-iterations were compared with the number of perceptual phenomena at stake, per edit.

Methodology and Reflections

“The main task of methodology”, in Bird’s reading of Peirce, is “the analysis of scientific discourse or argument” (Bird 1959 p.192). A methodology is, in this reading, taken to ensure that a method finds the proper and valid arguments that can answer the research question [or hypothesis] (Bird 1959). In this chapter, I will describe the procedure of inquiry for each study. Moreover, I will give an account of the research data that has been created, as well as motivate why it is appropriate to search for the answers to the research questions of this thesis within these data categories. The techniques for gathering data utilized will also be discussed, according to how well suited they are for generating the kind of valid or credible data needed in relation to the research questions. Possible disadvantages with the techniques will be addressed, as well as ethical aspects of the techniques used. Alternative techniques and procedures will be raised for consideration. I will also attend to the reliability or dependability of the results, especially the repeatability or confirmability of the studies, depending on whether the data is of a qualitative or quantitative nature. Within what framework would it be possible to repeat or confirm these studies with the desired accuracy or affirmation?

---

2 See appendix for table of perceptual phenomena.
The Literature Study on Film and TV Industry Crafts

The method chosen to identify film and TV industry crafts that create and process sound and image as design work (Paper A) is chosen primarily from the viewpoint of disinterest. Published literature is open and available for anyone who wants to repeat the analysis. The literature is written by experienced and recognized professionals, as well as published with established publishing houses. Thus, there is no reason to expect descriptions or propositions made in the works to be misleading or of marginal significance. There might be tendentious ideas on best practices, or preferred methods and routines. However, this has little impact on the core matters regarding industrial standard processes, or on what these processes must achieve, who is involved, in what way, and with what assignments. Being tendentious on these matters would undermine the author’s professionalism, for future collaboration with others in the audiovisual industry. On the contrary, a well-written work, recognized by others within the industry, generates extra credibility and thus strengthens the author’s professional position. Furthermore, published literature goes through a thorough script review process, where there are several re-iterations, and thoughts and formulations can be shaped to perfect the intended meaning.

After the search for appropriate literature, assortments were made of the found titles. A first grouping selected writings on production from an industrial perspective, selecting out literature regarding production as an activity for anyone interested. Secondly, the industrial perspective regarded production either as an executive/administrative activity, or as the actual manufacturing or processing of sound and moving images, or both. Only the second category was of interest in this study. The third grouping regarded whether the scope of the literature addressed limited and specific sub-sets of production, or covered all from ideation of sounds and images to the completed product, ready for distribution, with the latter as the category of interest. In the fourth step, literature was selected to cover a broad time span, from contemporary literature, to literature spanning back 40 years. Using this time span enabled the identification of recurring descriptions and propositions, as well as tendentious ideas to be omitted. The analysis focused on descriptions and propositions regarding ideation for, or execution of, activities that form sound or moving images, either through their creation or through their processing, such as recording, editing or adding effects. Any literature study of film and TV production using the same keywords for searching and the same procedure for selection of literature, as well as addressing the same categories of activities, will find similar data.

The alternative to a literature study would have been to interview comparable professionals from the film and TV industry. The immediate benefit of that method would have been to be able to use more directed questions, and to add supplementary questions, to further investigate specific subjects or issues. Since the literature study does not allow for that approach, it must instead be
more extensive to cover a variety of viewpoints. Nevertheless, some aspects might be missing. However, the aim of the literature study concerns a rather general level of understanding, so it is doubtful whether in-depth interview questions would have rendered data that would have generated radically different results than those from the literature analysis.

The immediate disadvantage of interviews, apart from time consumption, is that answers are limited to what comes to the interviewee’s mind during the limited time the interview lasts. There are also fewer opportunities in interviews for elaborating on thoughts and propositions, and no reviews or critical comments to utterances made. Neither must the interviewee regard the accuracy of utterances, from the viewpoint of industry peers evaluating them, since there is only indirect reference to a person, through the researcher. These drawbacks of the interview method call for a wide span of interviews to saturate the dataset, which makes the method time extensive. Therefore, I regard the literature study as the more tenable option, since it consumes less time, and addresses transferable and trustworthy data for common answers.

The Design Process Management Study

When intruding into the film and TV industry, the study on how new technologies made impact on production chains (Papers B and C, Swenberg 2012) exposed cruxed workflow situations in need of mending. After inventing MI-Flow as a tool for managers to foresee possible constraints in a production chain, the tool was evaluated regarding its purposefulness via an implemented prototype (cf. Blessing & Chakrabarti 2009). The core of a solution was developed, in co-operation with computer programmers and a compressionist, with the aim to secure the material quality of video files that are to be processed in a complex film industry production chain.

The tool was made accessible on the internet (www.mi-flow.se), and tested on independent film producers/managers (the target client group) for user assessment of its fundamental principles, by trying the tool for at least two alternative hypothetic production chains, and commenting on whether or not the tool could be useful in avoiding video file format and codec conflicts. The producers addressed were all male, reflecting the gender situation of the Swedish audiovisual industry at the time. Since we considered the producers as expert users, we only addressed a limited number (cf. Judge, Clarke, & Hawley 2011).

The data gathered from experts is expected to be rather dense, which then is expected to saturate the dataset already with relatively few participants. The data was gathered in two steps, where, first, a group interview around a “dummy” of the tool was discussed, and participants were paid for their time. Thereafter, other informants tried the fully developed tool for evaluation, and responded to us via electronic mail. Thus, the evaluation would not interfere
with the producers’ daily obligations, and they could do the testing whenever suitable for them.

Optional ways of carrying out this last evaluation would have been to visit the producers and observe them while testing, or gather a group of producers to test the tool in parallel at the same time, and afterwards, in either scenario, discuss the pros and cons of the tool. The benefit of these methods would have enabled the gathering of data about the actual use of the tool, its comprehensiveness, and get immediate response on details for development. Additionally, a dialog with the producers could have developed the data regarding the producers’ understanding of the tool, as well as on the need for such a tool.

Since we did not intend to develop the tool further, these efforts were considered too exhaustive, as we did not really need that much data. Other disadvantages of such methods are that it is rather difficult to gain access to film producers’ time, especially if they have to be at a particular place for a test session, and/or that discussions might steer thoughts in non spontaneous directions. In the end, any of the methods would render data that could be used to evaluate the tool. Therefore, our option was the more economic one.

The Study of a Film Editor’s Intentions and Viewers’ Perception

A more vast and exhaustive study was the study of a film editor’s intentions and viewers’ perception, where a three-step method is used in the full study, as well as in the pilot-study. These three steps consist of, first, observing and registering what the film editor does and intends while editing a film sequence. Thereafter, as a researcher, I intrude and make a new version of the film sequence, in which the film editor’s intentions with each edit are overturned. This intrusion is protocollled. Thirdly, the two versions are screened for participant viewers that are eye-tracked, and their gaze data is collected. As expected, there are improvements in the methods used from the pilot-study to the full study.

Involving a professional in a research project entails coming to terms with this individual regarding the conditions for participation, as well as the purpose of his/her participation. Two different film editors were engaged in the pilot-study, and the full study, respectively. Agreements were reached before both studies, but the film editor in the pilot-study was uneasy with the results from the study, and did not want to participate any further. Possible reasons for this unease was how I had introduced the purpose of the study to her, its analysis, or how the results were expected to be presented.

Consequently, I concluded that the involvement of another film editor required a more thorough introduction of the project to the editor, but also a deeper commitment on the part of the editor. Therefore, the film editor in the full study was invited to discuss the issues of hard-to-achieve fluid edits, i.e. typical joints between shots that resist becoming smooth cuts, in order to
outline what properties of moving images are common causes of editing constraints. Furthermore, the question was discussed of what data the eye tracking of viewers might give, and how this data might be interpreted in regard to the editor’s intentions. The film editor was aware of the possibility that viewers might not look at the film sequence the way she anticipated them to do.

Another aspect of involving a professional in a research project is that a professional is also an individual, and not all individuals of the same profession can be assumed to tackle all work tasks in the same way. What I have assumed regarding film editors is that they share a competence based on thinking, training, as well as experience that amounts to a knowledge of the alternatives in solving film-editing tasks, and are well acquainted with the methods required. Their professional credentials from decades in the trade are taken to ensure this competence, and they do not fail to meet the standards expected by the industry.

The idea is that the study could have been repeated with the engagement of any film editor with similar credentials. There is no absolute guarantee for meeting this condition, in any study. However, I found no better way of addressing the matter of recruiting an experienced professional either. The option would be to design the study differently, and to engage a number of individuals of the same profession either to come up with common solutions to the given tasks, or to give them the same tasks to solve, in parallel, and to compare the results. Either study design, while interesting, would have been very extensive.

An advantage in engaging a professional film editor was the input to the discussion on what film material to use, as well as her access to the documentary film material we came to adopt in the study. Having a dialogue on challenges that an audiovisual material can convey sharpens the focus on what material properties to seek, in order to put the editor to the test in trying to achieve fluid continuity. The outcome from this discussion was that the challenges of documentary film material seemed appropriate (cf. Audiovisual material). Whereas the editor’s acquaintance with the material, stemming from her making a full documentary film from that same material, was an asset in regard to the formulation of a story, since she knew from the start what could be narrated, or not, from that particular stock of footage.

Following the film editor during editing, either as observation or recording, serves the purpose of decoding editorial intentions regarding the film edits. Having access to the film editor while working gives the additional opportunity to intrude into the work process, or to interview afterwards, with or without elicitation of the recorded material. The pilot-study was mainly video-recorded. The editor preferred to work alone and discuss the work afterwards. The technique of continuously recording a computer screen, and the software in use – including sound and moving images – as a video file for later analysis, did not operate properly, and had to be abandoned. The assumed reason for the malfunction was shortage of computing power. For
the full study, a much more powerful computer was used, and the problem did not occur again. The tool utilized for these recordings was the software iShowU, which also recorded computer keystrokes and a monitoring web-camera mounted on the computer screen, viewing the film editor at work.

Thus, every moment of the editing process was captured for later analysis. Videos gave access to any iteration the editor made using the editing software, as well as her utterances and body movements. All these data categories can be used to see where in the editing process issues or problems occur, as well as when choices are made. Recordings ensure the possibility for repeated analysis. When the editing of a film sequence was completed, I interviewed each film editor about the intentions regarding the edits made, while watching the sequence (elicitation). During the full study, this interview was recorded. In the pilot-study it was not recorded, and only notes were taken. The substantial difference was that when the elicitation was recorded, the links to the audio-visual material on the screen were preserved and were thus more apparent for the analysis.

Knowing the intentions of the film editor for each edit of a film sequence, it was possible for me to make a new version of each sequence that I manipulated so that these intentions were countered by the adding or removal of frames at each edit point. This was done for each edit in the pilot-study, as well as in the full study. In each study, the film sequence in question was adopted as stimuli in the later eye-tracking part of the study. Viewers had both versions screened in random order, as a within-subjects set-up, while their eyes were recorded, and pupil reflections measured. It was intended to possibly verify or falsify the editor’s intentions by comparing the recorded eye data for each version of the film sequence.

The possible disadvantage with this method is that viewers recognize a film they have just seen, and thus watch differently the second time. This phenomenon is addressed by randomizing the viewing order of the sequences, and engaging enough participants. The alternative would have been to engage more participants who only watched one version, as a between-subjects set-up. However, that would have led to differences in viewing patterns and pupil sizes that individuals exhibit.

The eye tracking of viewers was set up rather differently in the pilot-study and the full study. First, the recruitment of participant viewers differed. In the pilot-study, I wanted a broad scope of viewers to even out differences of age, gender, educational level, or film/image knowledge. In the full study, the group of participants was limited to college students, which conditions the average age and educational level. The study was still gender balanced, and also collected the film/image acknowledge of the participants. Eye behavior is understood to be rather consistent for adults, after reaching the teen ages (Weale 1991, Leat, Yadav, & Irving 2009, Doron, Spierer, & Polat 2015), therefore, the limitation in age range in the full study is assumed not to matter.
Educational level is accounted for in Paper G, and supposedly delimits the validity of the results primarily to people with formal education, spanning from initiated to completed bachelor’s studies, since education supposedly influences viewing patterns. Many earlier eye-tracking studies are subordinate to similar conditions, and the effect of this limitation is that results therefore are considered as indicative or hypothesis-rendering (Holmqvist et al. 2011 p.87). In the full study, each experiment run lasted between 20 and 30 minutes (depending, for example, on how much time participants spent on answering questions).

Eye-tracking procedure includes three steps (cf. Holmqvist et al. 2011): First, one must consider participants’ conditions while partaking in the experiment, which might differ between individuals on their arrival at the lab. They must therefore be given some time to rest and settle, or become alert, in order to reach a psychological state as uniform as possible, when the actual eye tracking is conducted. This step was ignored in the pilot-study, due to inexperience on my behalf. During the full study, all participants were given at least 15 minutes to sit down, have a biscuit (to even out blood sugar levels) with a cup of tea or coffee, or a glass of mineral water, before the eye tracking started.

Participants leaving the lab were not allowed to meet the entering participants, and were also asked not to talk about the experiment until after the recording days. The reason for this procedure is that pre-experiment knowledge about the purpose of the test supposedly has effect on behavior during the experiment (Holmqvist et al. 2011). It is known that tasks during eye tracking affect the viewing (Yarbus 1967), and that knowledge about experiment purposes are interpreted into tasks by participants (Holmqvist et al. 2011 p.77-79). Therefore, participants were told as little as possible about the experiments beforehand, but were promised, and given, full information afterwards (cf. Holmqvist et al. 2011). In the pilot-study, the information was given orally. During the full study, the participants were given the information in writing, before and after the experiment, so as not to accidently omit any information, or place oral emphasis on various parts of the information.

The participants were then given the chance to reject the usage of their eye data for further research analysis, or to give their written consent. On rejection, the eye data was immediately deleted from the eye-tracking computer while the participant looked on.
Figure 28. Distances between: participants and screen, speakers, and screen and speakers. Positions $X_1$ as well as $X_2$, are measure points for sound and light emission. $X_2$ is also the participants’ viewing position.

Figure 29. A: The researcher’s position (chair + table with computer, screen, mouse, keyboard, and sound controls). B: participant’s position (chair + table + stimuli screen, mouse [M], keyboard [K], and speakers [S]). Speakers were placed on a separate table to avoid the spread of vibrations to the eye-tracking camera. The positions of the researcher and the participant, respectively, are separated by a screen wall to shield the participant from visual disturbance from the researcher’s activities.
The run of the experiment consisted of having each participant, one at the
time, seated before a computer screen, with the eyes at approximately 70 cm
distance from it (see Figure 28). A stationary eye-tracker (SMI RED60*/
RED250** ran with SMI iView software) was mounted under the screen, and
a couple of supplementary loudspeakers stood just behind the screen on either
side (see Figure 29).

All participants had normal, or corrected-to-normal eyesight. All infor-
mation to participants during the run of the experiment was given on the
screen, as text. The experimenter sat behind a screen wall**, or behind the
participants*, so as not to disturb the participants while managing the expe-
iment. Each experiment (run with SMI Experiment Center software) includes
an initial calibration of eye data capturing, and a validation of the calibration.
According to eye-tracking research practice, the validation deviation should
not exceed 1° of visual angle when gaze data on information search is con-
ducted (cf. Holmqvist et al. 2011), a limit that was retained during this study.
Thereafter, information for each different step of the experiment is presented,
as well as leading questions to be answered after each stimulus. These ques-
tions function as tasks for the participants, and are formulated so as to pro-
mote a natural viewing of the film sequences, e.g.: “What genre would you
suggest that this film belongs to?”. The participant is then given optional film
genre alternatives to answer after viewing. This data is not considered during
the data analysis, and this is explained to the participants in the post-exper-
iment information. In total, I used at least 45 minutes of each participant’s
time in the full study. The participants were rewarded with a cinema ticket
each, worth approximately €10.

The full study was also conducted by means of an experiment that was run
interwoven with stimuli from another study, conducted by my fellow doctoral
student, Per Erik Eriksson. Researcher co-operation in collecting data en-
hances data quality and credibility (Elliott et al. 1999). The pilot-study ran
only the short 37-second film sequence in use, twice. Since I screened two
versions of the same film sequence (with differences not noticeable to the par-
ticipants), there is an advantage of interweaving stimuli from different studies
in the experiment, so that the participants got to see something else between
the two versions, and were thus at least somewhat less subjected to repetition.
In addition, some participants expressed confusion over the purpose of the
study, after having a number of different stimuli screened before them. Again,
this is a disadvantage, if participants actually behave strangely, due to these
concerns. If this is not the case, it is beneficial for the study, since knowing
the purpose of a study in advance may affect participants’ viewing of the film
sequences.

Regarding eye data, I constructed my own rough categories of eye behavior
to search for in the pilot-study analysis, consisting of resting, consequent
movements, and radical movements, which were less well-defined than the
ones traditionally used in the studies of film perception: fixations, saccades
and pursuits (d’Ydewalle, Desmet, & Van Rensbergen 1998, Smith 2005, Carmi 2007). However, in the full study, I focused on three well-established categories: Areas of interest (AoI:s) appointed by the editor, Saccades, and pupil dilations after edit points. These analytical categories were extracted by means of SMI BeGaze software. In terms of proper measures for the estimated effects, a more thorough definition of the categories used in the pilot-study would have provided a well-functioning measurement also in the full study. However, to match previous research, and be able to compare results, I used the measures established in the literature (see section Operationalization). This will make my results comparable to future research as well. An additional, optional, measure would have been micro-saccades after edits, but this category was abandoned due to time shortage for the eye data analysis.

During the analysis of the full study, hits (eyes fixating at least once) within Areas-of-Interest (AoI:s) were counted for both versions, in the time span 120-400ms after an edit. Comparisons were made for each sequence. The amount of saccades in the same time span was measured for each version of the sequence. The proportion of pupil dilations (compared to pupil sizes during insignificant visual input), which was 200-500ms after edits, was measured, for each version of the film sequence. Differences were tested statistically by T-tests, or Binominal distribution.

The alternative methods of capturing viewers’ perception of film edits drawn from previous research are predominantly of an inductive kind. Such studies either test viewers’ visual responses to graphical features, occurring, moving and disappearing from a screen (e.g. Williams & Simons 2000, Smith 2005), or record manual responses to edits (button-pushing), (e.g. Geiger & Reeves 1993, Tam et al. 1995, Shimamura et al. 2015), or estimate viewers’ preferences for different edit points (e.g. Hecht & Kalkofen 2009). Either method gives results that must be superimposed upon actual film viewing, as an argument of similarity, which might not stand the test against recording actual eye behavior during film viewing (see Discussion and Conclusion).
Results

The scientific research results presented are categorizations and conceptualizations made from the assortment taking place during data analysis, along with given reasons for the concepts and categories used. The results from the studies underpinning this thesis are described in this chapter. How the results answer the research questions will be further attended to in the Discussion and Conclusion chapter, below.

Firstly, the principal results from the studies are presented in the order they are addressed in the analysis (steps 1-5): the relevance of studying film and TV production as audiovisual design, the reasons for studying digital audiovisual material treatment, the film editor’s employment of perception, measurements of perceptual precision in film editing, and the film editor’s intentions regarding viewers’ gazes.

Thereafter, the detailed results are presented to the reader interested in their configuration, according to how they support the principal results. The results are credible, primarily, in relation to the contexts from where the data is derived. However, I will suggest further implications of the results in the Discussion and Conclusion chapter.

Principal results

Results can be formulated in detail, or in summary. In this section, the results will be presented according to the result category of 1-5. They will be summarized and presented from the viewpoint of their key contributions to answering the research questions. The first, second, third, and fourth results, in combination, will address the influence of perceptual precision on film editing, while the third, fourth, and fifth results, together, relate perceptual precision to design intentions. The results are presented in detail and at length below.
Result 1:

Legitimacy for studying film and TV production as design

The literature analysis of production literature, where experienced and renowned film and TV practitioners reason about, and state, their views on what is crucial in film and TV production (Paper A) suggests that the creation and processing of sound and moving images in film and TV production can benefit from being studied as instances of audiovisual design.

The results of this analysis are that the work executed by crafts-people handling sound and image share similar properties with other design work: it constantly engages in (aesthetic) problem-solving and (communicative) meaning-making, two aspects that are very often intertwined. Furthermore, work processes, as well as problems and issues that engage practitioners of different crafts are of a similar nature to design processes, design problems and issues corresponding to those studied within design research.

Moreover, if design is the ideation and shaping of an artefact, and production is the realization of that same artifact, the literature analysis suggests little trade-off in trying to distinguish between design and production as actual, distinctly different activities in audiovisual communication areas, such as film and TV production. The craft practitioners’ statements reveal that most design work takes place during the production phases. Correspondingly, when the design is completed most of the production is too. Thereby, the division between, on the one hand, the design, and, on the other, the implementation of a design, is blurred beyond any useful clarity (see Figure 30). These results support the starting point for addressing film editing as a kind of audiovisual design work.

![Design strata](image)

**Figure 30.** The different activities of design and production largely overlap within the film and TV industry. Design largely takes place during production, or vice versa. (Adapted from Swenberg & Eriksson, *Forthcoming*)
Result 2:
The material’s effect on creative space

Employing a design process management perspective, the study of film and TV production as design processes contextualizes certain crafts and work roles within, where the importance of each design agent’s material creative space is emphasized (Swenberg 2012, and Paper B). Such a creative space consists of the dimensions time for design work, tool capacity and user skills, and the material being processed with its properties (see Figure 31). Maintaining these creative spaces in complex design processes in the audiovisual field by minimizing time waste, and keeping proper material quality, are critical success factors for each agent’s design work, as well as for the collective of designers within a wider and more complex design process.

![Diagram](attachment://diagram.png)

**Figure 31.** Dimensions of a designer’s creative space in moving image production. (Figure and caption as in Swenberg 2012)

In this study, the use of a management tool to avoid digital material conflicts was tested, and confirmed to provide a much-needed support for audiovisual industry managers (Paper O). Here, the most important of these results regards how the processing of digital materials consumes a designer’s time, which otherwise could be spent more creatively. Effective handling of digital material is the core to maintaining the design agent’s creative space, and to target creativity to where it is most desired.
Result 3:
Use of perception in film editing

Perception affects creativity. The analysis of the observational recordings shows that the film editor employs perception, and that it has bi-directional implications on, and for, the designing of particular edits, editing principles, as well as semiotics of concrete details (see Figure 32).

Figure 32. PIFEM. Perceptual Implications in Film Editing Model: perceptual implications according to the gathered data, data type per theoretical category. Arrows indicate direction of influence between categories recognized in this study. The design category is treated as overlapping semiotics and editing principles.

Further, the results from the video analysis, together with the analysis of the completed film sequence, show how a film editor considers perceptual aspects in the editing process, both audio and visual, and that perceptual phenomena are considered by the film editor for almost all edits (Paper F). There is also a correlation between, on the one hand, how many such phenomena there are at stake at an edit point, and, on the other hand, how long it takes for the editor to complete the edit to satisfaction, as well as how many iterations (reprocessing) of the edit it takes.
Result 4:

Viewers’ perceptual reactions to film edits

Film viewers’ (audio-)visual perception of film edits is addressed, when eye tracking viewers watching the film editor’s documentary film sequence (Condition 1), https://vimeo.com/210944884/be8001e786, as well as a sequence where her intentions with each edit point were counteracted (Condition 2), https://vimeo.com/210953506/9ed3cb3b6e. A majority of viewers look where the editor intended, in both versions, so this editorial direction is insensitive to frame-matching. Saccades and pupil dilation were measured to distinguish between viewers’ perceptual reactions to the smoother edits made by the film editor (Condition 1), and to the harsher ones I had manipulated (Condition 2). Such a slight shift of the edit point, as up to six frames, causes a 29% increase in viewers’ eye movements after an edit point (Paper G). The viewers’ pupils also expand more after these manipulated edits (Condition 2), than they do after the editor’s version of the sequence (Condition 1), which indicates a higher cognitive load for the second condition (Paper G). The importance of perceptual precision in the audiovisual design of film edits is thereby indicated.

Result 5:

Editing intentions in film editing

The first result from the study of a film editor’s intentions and viewers’ perception was a developed mixed method, VOSMET (Video Observations and Screen-recordings plus Manipulation and Eye Tracking). This method enables the study of visual intentions in film editing in relation to viewers’ visual perception of film edits (Papers D and E). The VOSMET method can also evaluate whether film edits are smooth or harsh for the viewer.

Secondly, the film editor’s intentions are captured, both in her comments during work, and during elicitation. She intends her viewers to look at specific areas within the image frame, but not for all edits. She also intends appointed audiovisual phenomena to either capture viewers’ attention, or not. It is clear that the film editor strives for continuity for most edits, but that she wants discontinuity for others.
A short cross-reading of the results

The results from these studies together indicate the importance of perceptual precision in (audio-)visual design. When viewers are not directed perceptually, they start searching for the most important visual information. Then, they either look at less informative areas of an image, adjust their gaze, utilize more brain capacity, while looking at the image, or use a combination of any of these three. The direction of the gaze is achieved by the designer, framing the viewer’s interest regarding where to look. This framing consists of removing non-important (audio-)visual transients (details that stand out from the context) that occur as competitors to the intended information.

The current results will be analyzed in relation to how the handling of a material consumes time, and thus impinges on the creative space of a designer. Perception is engaged in handling competing audiovisual material details that are all likely to attract attention and designers are challenged in directing attention. Perceptual precision seems to involve control of the less important competing details. The implications of these results on design processes will be discussed further below.
Detailed results

The results summarized above will be considered in detail in this section, and presented below at length, per category.

Film and TV production addressed as design

Searching in major library and research databases for extensive writings by prominent and experienced craftsmen in the field of film and TV production industry resulted in books containing the following terms in their titles: “film production”, “TV production”, “television production”, “moving image production”, or “motion picture production”. These publications were analyzed for how they problematize and discuss questions regarding the executing of sound and image work, in relation to design theory. The underpinning results for considering film and TV production as audiovisual design are four.

First, working with the notions that design is the ideation and shaping of an artefact, and production is the realization of that same artifact, there is an unclear distinction, on the one hand, between the phases of film and TV production processes that can be recognized as design, where sound and image are created and processed, and, on the other hand, the phases that can be distinguished as mere production (Paper A). Audiovisual communication is multimodal, consisting of interwoven forms of expressions that continuously differ in constitution as time runs. In order to reach the desired expression-in-combination of sound and moving image, it is necessary to try out different expressions, and to iterate their combinations, until a satisfactory result is reached. This is typical of efforts internal to any design process. The step from the completion of a design to production and a complete product is very small. Commonly, there are only small technical adjustments to the product necessary for it to be completed. This is the step between a design phase and the phase of implementation of that design. The production process has largely taken place during the design process. In other words, the design process of sound and moving image (the material) and its production process is essentially the same process within the film and TV production industry.

Second, film and TV production regularly make use of the established design principles of (aesthetic) problem-solving, or (audiovisual) meaning-making. Individuals contribute to collective efforts in developing, adapting, adjusting and altering audiovisual stories into comprehensive and enjoyable audiovisual objects that become additions to the existing stock of such objects, which are recognizable as significant features of design work. The comprehension that a film and TV production team seeks for its audience could well be understood as design cognition. Any sound or image creating, or processing craft activity, shapes audiovisual material as acts that create meaning. Each such activity requires that the individual understands the production
situation rationally, to know what design principle to use, what action to take, and to realize how to execute the proper activity.

The third result is that the production processes within the film and TV industry resemble other design processes and can benefit from being studied from a design processes research perspective. Negotiations of meaning, function and appearance, regarding the material qualities and complexities of sound and moving image expressions, are essential parts of film and TV production processes, as similar issues are in design processes. Another similarity is testing expressions on an audience for assessment, to reduce the uncertainty some estimations must be made under. The uses of aesthetic conventions or innovations, as well as design principles, are characteristics of both fields. In addition, as for design processes, film and TV production processes are also subjected to the formation of sub-processes, for example: script development, pre-visualization, recording, editing, sound design and titling. Further, process and method optimization for process efficiency is employed, which are recurring topics in design research, considered for distributing bottle-neck resources, human or mechanical. When new production methods are employed in film and TV production, new processes must be developed to match the old and new methods to each other. Of special importance is the identification of a process “drivers” – such as craftspeople’s creativity, and “constraints” – e.g. risk of time-loss from technical hassles. These are usual topics in design process research. Other design process characteristics studied, resembling the situation within the film and TV industry, are creativity, complexities of communication, and technical constraints issues.

Fourthly, the result regarding issues and problems that occur during the creation and processing of sound and moving images in film and TV production is that these are often of the same nature as design issues and problems studied within design research. Issues and problems need consideration from the viewpoint of “bounded rationality”, i.e. treated on their own terms, according to the needs of a situation, beyond conventional solutions. Concrete aesthetic considerations impact the meaning of an audiovisual product, such as aspects like size, color, shape, texture, and meaning affordances must be related to story and genre, for instance. Emotional impact is often desired even for non-entertainment productions, aspects of film and TV production that recurrently have to be considered when addressing an audience and a context. The compilations of sound and image components must be synthesized into solutions that implement intentions and ideas coherently. From these conditions, appropriate tools and methods must be chosen in order to achieve the expected expressions. Often, such issues and problems amount to rather complex situations, which resemble the kinds of complexities studied within design research.
Digital materials and constraints on creative space

The detailed results from the study of audiovisual industry organization as design process management are as follows: it is important to maintain each design agent’s creative space, in such rather multifaceted processes as film and TV production processes (Paper B). This intricacy calls for attention to a bottleneck issue that may have severe impact on the processes, if not treated purposefully: a design agent’s creative space configures and affects others’ in a production chain consisting of several design agents (Swenberg 2012). Each design agent has a certain dedicated time span for doing his/her work tasks, executed by means of software tools on digital material (Paper B). These three material factors are linked: Digital material can be configured in many different ways, by means of codecs and file formats. Tools differ in their sophistication, complexity and capacity. In this regard, two parameters are their capacity to handle and transform (or convert) digital materials. Usually there is one, or just a few, digital materials that are well suited for each tool. If a digital material arrives at a design agent’s work-station in a non-ideal digital configuration (assumed high tool capacity), the design agent will have to use time and tool capacity to transform the material into a better suited configuration. For large video files, such transformations consume significant proportions of time. Alternatively, the design agent will have to settle for a poorer material that cannot be processed as much as desired. In any of the two cases, the creative space shrinks, either through time-loss, or a lesser degree of possible processing. A third case regards low tool capacity for any kind of digital material, or an agent’s incomplete mastering of the tool, which will obviously limit the space for creativity (Paper B). Taken together, tool capacity and user competence are termed “expressive potential” in this context. These three material factors, time, digital material, and expressive potential affect each other (see Figure 31).

The complex system with many design agents is sensitive to how well each agent functions, and to how well the design work flows through the system (Swenberg 2012). In the case of film and TV industry workflows, choosing the proper tools for each design activity and avoiding transformations of materials are key issues in order not to waste time (Paper B). To maintain a well-functioning design workflow – that is, maintained creative spaces for each design agent – throughout the system, management is crucial, and management tools can be supportive in this regard (Paper C). In an attempt to suggest mending for the audiovisual industrial situation, the digital check-list MI-Flow was developed (www.mi-flow.se) as a design management support tool, and evaluated against industry representatives (Paper C). Its function, which was confirmed as effectual, was to reveal conflicts in production chains regarding the digital configuration of materials (video file formats and codecs). Revealing conflicting digital materials would help in sustaining the creative space of each design agent in the production chain, by ensuring the proper
choice of materials according to chosen tools. The overall aim is always to be able to spend as much time as necessary on the creative adaption of sound and image expressions, with an adequate tool, and proper digital material.

The film editor’s use of perception when editing film

One aspect of the processing of digital sound and image material is to regard its perceptual qualities, to estimate what sound and image features trigger human attention, and in what way, as well as what features fail to attract attention. The results from the study of a film editor’s intentions and viewers’ perception contain the film editor’s expressions for the challenges of each edit, including material challenges: “These shots are very similar in framing. Quite. And the cut is between a very dark and a very light shot.” These conditions make visual transients apparent. Perception is also often mentioned as a key challenge in achieving a satisfactory edit: “When a cut flashes: there is something happening at that edit that disrupts, and which gets one out of the story, or out of…. …this flow.” The analyses of the videos also show that the film editor’s employment of perception has mutual implications between, on the one hand, perception, and, on the other, the editing as design work, editing principles exploited, and semiotics in audiovisual details (see Figure 32). This three-fold involvement of perceptual aspects during film editing regards the treatment of the audiovisual material. Perception is involved in the choice of frame-to-frame matching of edits, the removal or replacing of sounds, the altering of frame rate, and decisions on how much movement there should be within the image frame, before and after the edit points. All these are design aspects. The motivation for these design actions is occasionally to remove sound or image features which draw attention to themselves, and away from the intended audiovisual features for attention. At other times, such features are estimated by the editor as details that might cause unwanted interpretation of meaning. Moreover, at the instances of edit points, the interplay of audiovisual features on each side of the edit constitutes whether the edit belongs to the respective category of continuity or discontinuity edits, which adhere to different editing principles.

When I examine how the film editor deals with perceptual phenomena, I find relations between perceptual phenomena at stake for each edit, time consumption, and number of iterations required for the film editor to complete the edit to her satisfaction (Paper F). When smooth (continuity) edits are desired, the editor avoids any (audio-)visual transient that might attract the viewer’s attention, by either removing it or avoiding it through choice of frame-matching (omitting frames with visual transients). It happens that the editor alters the running speed of the moving images and freezes frames in order to evade movements that draw attention to an edit. There are, on average, eight (8) perceptual phenomena at stake for each of the 20 edits in the examined 3’10” documentary film sequence, spanning from one (1) to 14 phenomena per
edit. These phenomena are compared to the time needed by the film editor to complete each edit in the sequence to satisfaction, and to the number of occasions the editor re-worked the edits until satisfied. The results show strong correlations, r=0.58 (p=0.008) for the time relation, and r=0.53 (p=0.017) for the re-workings relation. So, the more perceptual phenomena at stake for a film edit, the more time it takes for the editor to complete it to satisfaction, and the more re-workings it requires as well. The results also show that the film editor is well aware of her attention towards her own perception when she assesses the edits during the editing process: “I am attentive towards my own gaze, and trust that everyone watches [the image] as I do”, she says (Paper F).

Viewers’ gaze responses to perceptual precision in film edits

How film viewers respond to the perceptual lead of a film editor is the next query. The set of results from the eye tracking of viewers watching the completed documentary film sequence reveals three things: (1) whether or not film viewers did, or did not, look at the film across the film edits, in the way the film editor intended; (2) whether viewers’ eye movements are sensitive to the film editor’s perceptual precision regarding the film edits; and (3) whether the perceptual precision matters to the viewers’ cognitive load during viewing (Paper G). Pupil size is a measure of cognitive load, whereas amount of eye movement corresponds to the extent of visual search conducted. When edits provoke more search or more cognitive load, viewers ultimately experience fatigue viewing such sequences. To be able to compare the effect of the perceptual precision the editor had achieved for each film edit (the editor’s version is Condition 1, https://vimeo.com/210944884/be8801e786), I created a new version of the film sequence where the edits were all altered between one and three frames (Condition 2, https://vimeo.com/210953506/9ed3cb3b6e). My aim was to try to undo the intention the editor had expressed regarding each one of the edits. The two versions of the documentary film sequence were screened for 33 viewers, one at a time, before a stationary eye-tracker (SMI RED250).

The eye data was analyzed for gaze hits in the Areas of Interest (AoIs) that the editor had pointed out, and for eye movements, in a short time window (120-400 ms) after the edit points, as well as for pupil sizes (200-500ms after edit points), where time windows correspond to perceptual response to the edits. Regarding average hits in AoIs, there were 54,0% (SD=28,1) hits for the editor’s version (Condition 1), and 58,5% (SD=29,9) hits for mine (Condition 2), for first viewings, and the small difference was not significant (p=0,095 from T-test), (Figures 33 and 34, and Paper G, Table 2). The difference in frame-matching had no detectable effect.
Figure 33. Diagram of average hit ratios in Areas-of-Interest (AoS) for Conditions 1 and 2 (two versions of the same documentary film sequence) when eye tracking film viewers. A hit is when a viewer’s eyes have fixated within an AoI. Standard deviation (SD) for each condition is indicated.

Figure 34. Example of an Area-of-Interest (AoI) appointed by the film editor as an intended target area for film viewers’ gazes (here, after edit #18 in the documentary film sequence). Eye data statistics are provided as Key Performance Indicators in the box on the right corner: the hit ratio for this AoI (Condition 1) was 45.5%.

Saccadic eye movements showed a significant increase in number by 29.4% (p<0.0003; pairwise difference), (see Figure 35). Pupil dilations also differed. When pupil sizes were compared to an eye at ease, looking at a grey screen, requiring no effort, the edits in the editor’s sequence (Condition 1) rendered an increase in average pupil size by 7.9% for all viewers as a group, which equals an “intermediate” cognitive task level. When the same group of viewers watched my version (Condition 2), the average pupil size after edits increased by 10.1%, which, in turn, equals cognitive tasks of high difficulty (see Figure 35, Paper G, and Table 3). The difference has a strength of p=0.023.
Viewers’ perceptual response to events occurring in the film can also be detected by means of pupil dilation as an indication of event detection processing. I tested viewer responses towards coarse events as well as fine-grained events. Regarding coarse events, there was an increase in average pupil size after edits in both versions of the film sequence: 10.6% increase for the editor’s version (Condition 1), and 11.3% increase for mine (Condition 2). The difference was not significant though (p=0.31), (cf. Figure 35). For fine-grained events, there was also a difference, which almost reached significance: the editor’s version (Condition 1) showed a 10.1% increase in average pupil size, and my version (Condition 2) showed an 11.3% increase (p=0.056), (cf. Figure 35). These results all match a high level of cognitive load, which is to be expected for detecting starts and ends of new events. The pupil size measure, though, indicates amplified cognitive load for viewers when film edits are not as smooth as possible, compared to when they are. There is also reason to expect that the viewer’s experience of how filmic events unfold before the eye, and are absorbed into the mind, is affected negatively by harsh edits, since the perception of fine-grained events might be disturbed.

**Figure 35.** Comparative diagram regarding saccade frequency, and pupil dilation (PD) for Conditions 1 and 2 when film viewer’s eyes have been tracked. Saccade frequency is measured after edit points, and pupil dilation after edit points, as well as related to edits as visual events.
The film editor’s intentions when editing film

The results from the study of a film editor’s intentions and viewers’ perception contain a set that regards the editor’s intentions and use of perception. Prior to obtaining this set of results, the method developed for studying visual intentions in film editing in relation to viewers’ visual perception is a result in its own right (Papers D and E). The method, VOSMET (Video Observations and Screen-recordings plus Manipulation and Eye Tracking), manages to capture several aspects of data in the same video recordings: (1) the body language and utterances of both the film editor during editing, and the observing researcher (me) sitting next to her, as well as (2) occurring conversation; (3) the screen with all data and activities that take place there, and (4) all keystrokes executed on the computer used for editing. These recordings are also suitable for (5) video elicitations after the work shift or complete work is finished. This combination of data makes access to intentions and perceptual considerations available. The method also includes a stage of eye tracking viewers, by means of which (6) the fulfillment of intentions can be assessed, and (7) the nature of edits – smooth or harsh to viewers’ perception – can be determined. The eye tracking method delivers data that shows viewers’ pupil sizes, amounts of eye movements, and viewing of areas-of-interest.

The results from the analysis of the video recordings, of both editing work and elicitation, show that the film editor has intentions for viewers’ attention and for areas of the image to be attended to, for many, but not all, edits (see Table 2). “I reckon that one is somewhere on this blue coloration [with one’s gaze]. But I don’t know. Yet, that doesn’t matter, since he [the protagonist] grasps the gaza with his hand.” Specific (audio-)visual phenomena are pointed out as supposedly capturing the viewer’s attention, which sometimes is desired, and sometimes not: “This one [edit] I do not get along with. […] There is this little blink.” Most often, the film editor strives for the smoother continuity edits, but occasionally also discontinuity edits, when the story takes a leap forward.

| Table 2. Continuity or discontinuity intentions for edits. |
|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Edit #     | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Continuity | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | N | N | N | Y | Y | N | N | Y | Y | Y |

Y = yes; N = no (= discontinuity). Editorial intentions regarding where within the image frame viewers are supposed to look are linked with continuity edits only.
Addressing the results

How moving image edits can be concealed before a trained eye, something which has intrigued me from the outset of my research, is perhaps not fully revealed by these results. It is becoming more apparent, though, that precision regarding how edits are designed from a perceptual stance is important, both from the editor’s, and from the audience’s point-of-view.

The issue addressed by my first research question, *What does perceptual precision attain in the design of continuity edits for viewers’ film decoding?* finds its answer in a combination of all the results, perhaps with the exclusion of the last results, regarding editorial intentions.

The second research question, *In what ways could (audio-)visual intentions be dependent on perceptual precision in the design of continuity edits?* is addressed in the results regarding the film editor’s intentions when editing film, and viewers’ gaze responses to perceptual precision in film edits.

How the research questions are answered by the results will be discussed next, in the *Analysis* chapter.
Analysis

An analysis is here seen as engaged in addressing research results in relation to research questions and relevant theory. Thus, an analysis contextualizes the results academically, by making connections between current results and results from previous research, and proposing how to understand the results. Further implications and consequences of the analysis will remain to the following discussion.

This analysis chapter will present the analytical approaches used in addressing the research questions of the thesis. The research questions are considered one at the time, with a breakdown into inherent aspects. For each aspect, there is a presentation of what results are appropriate to consider, as well as by which model(-s) of theories-in-interplay the results are analyzed. The discussion will thereafter be conducted from these analyses.

Analysis addressing the first Research Question

I will now address the first research question, which is of a contextualizing character: *What does perceptual precision attain in the design of continuity edits for viewers’ film decoding?*

This question contains several aspects that need observance:

1. film editing is assumed to be design, yet, that assumption needs justification;
2. prerequisites for the design of film edits regard the material handling of shots and sounds, so the audiovisual material’s conditioning of the design must be considered together with its handling;
3. the designer’s care for perceptual aspects must be addressed;
4. viewers’ film decoding must be considered, and related to their perception.

These four aspects of the research question will be addressed in different parts of the analysis. The notion of (a designer’s) perceptual precision will be related to viewers’ perception in the discussion, below.
Analysis approach and results considered

The analyses regarding the first research question will apply some of the theories in interplay, presented in the first part of the section Model Application and Interplay. Firstly, when addressing film editing as design (1) I will turn to the first result, and use the Design Profession Interplay Model regarding film editing as a design profession (see Figure 20). These theories contribute to the understanding of film editing activities from a “within”-perspective, as film editors understand them, as well as from an academic perspective, where audiovisual knowledge can be unveiled, and eventually brought to public knowledge.

During the next step in the analysis, the second result will be examined by means of some theories in interplay, according to the FEAD-model, that regards the handling of materials (2), in design, crafts, and film and TV production (see Figure 21). These theories take into account design as problem-solving, as well as meaning-making, as design problems are addressed and design decisions are made. Aesthetic problem-solving in processing a documentary material is linked with creativity and digital materiality as in the adapted Creative Space model in Figure 36.

Figure 36. The Creative Space model adapted (from Swenberg 2012) to include any material property (not just digital code).

The third step of the analysis deals with the third result, considering documentary film editing work and especially whether perceptual phenomena are dealt with (3), using the Design Profession Interplay Model and the FEAD-model (Figures 20 and 21). Dealing with perceptual phenomena can be addressed as a characteristic activity, or as the handling of material in problem-solving, as well as in meaning-making while making an artefact.
The fourth step of the analysis regards the fourth result, viewers’ gazes when watching two versions of the documentary film sequence (4): the original one made by the editor and an altered one where I had manipulated the edits. Theories in interplay according to the FIPA-model are found in Figure 22. A difference in gaze behavior would reveal whether or not perceptual precision during editing matters. Gazes were analyzed for hits in the editor’s intended AoI:s, for saccade frequency, and for pupil dilation, in time windows that correspond to established reaction times after an edit point (cf. Smith 2005, Seeber 2013).

Analysis – step 1: film editing as design

The first step of the analysis regards the justification of considering film editing as a design work. The first result, stemming from the literature study, concludes that work with handling audiovisual material during film and TV production is better understood as design work, since there is an almost complete overlap between design and production activities. Hence, considering any handling of audiovisual material design work will lend the work to be studied from an established design research perspective. Film editing is then included in this extensive category of work, by default, since handling and processing of sound and moving image is what its work tasks consist of.

How film editing, then, can benefit from being regarded as a design work becomes clear when we use the Design Profession Interplay Model (Figure 20). The film editing writings of experienced craftspeople contain the self-understanding of the trade, including its typical work tasks. It will be possible to

---

Figure 20. The Design Profession Interplay Model: design professions can be understood from different perspectives: from within the community of the profession; from academic interest in professions; and from public agencies. From all three perspectives, understanding activities is needed, and makes a central contribution to how a profession is distinguished. Film editing can be understood as design from all three perspectives, considering its activities, where trimming edits with perceptual precision could be one.
further explore the film editing trade by dividing such understanding into categories that recognize knowledge common to the trade, typical actions performed, as well as what qualities a film editor needs to possess. The use of perceptual precision when trimming edits is candidate to be such a kind of knowledge, employed during a typical activity. Such understanding aligns with a from-within perspective of design professions advocated by Adams et al. (2011), recognizing knowing, acting, and being, as main categories.

From an academic point of view, there is a gap in knowledge regarding film editing which can be addressed if it is treated as a design craft. Film editing aspects, such as emotion – e.g. rhythm; and experience, for instance perception, can thus be studied and better understood from the point of being part of recurring activities. Studying crafts by activities, in general, is also suggested by Niedderer and Townsend (2014) as necessary to deepen our understanding of crafts.

These approaches to understanding film editing as a design profession – from within, as well as from an academic point of interest – have the potential to deliver distinguishable characteristics of the trade. Such characteristics can be used to formulate its categories of work into a professional topography, by which it can be publicly recognized. This means pointing out and delineating the most typical activities performed by film editors, in a manner that is accessible to the public. By a design recognition, the reticence which, according to Murch (2006), film editing has maintained since its emergence as a profession, can be overcome. The need for such recognition for design professions is known and its mending encouraged (Smith & Whitfield 2005).

Analysis – step 2: handling of audiovisual material

The second step of the analysis considers result #2: a material’s effect on the creative space of a designer within the field of moving image production, especially the processing time parameter. It is preferred to target time and creativity to where it is most desired, i.e. where it contributes most to the artistic expression. In film editing, the artistic expression resides in properties regarding how the sound and moving image flow as a running stream of audiovisuality.

However, with reference to the FEAD-model of Figure 21, the handling of material in documentary film editing is, in my study, as in the usual case, conditioned by not being shot with the editing process in consideration. Hence, audiovisual matching involves many unpredicted challenges: perceptual phenomena, such as noise, or visual transients occur, semiotic properties of the material might be contradictive, and, thus, design problems become more complex, and design decisions harder to reach. Nevertheless, the possibility of making many tries and versions until a satisfactory result is accomplished prevails, when the material is digital. This double-foldness of digital video material is well known (cf. Manovich 2001 & 2012), as is the sensitivity
to credibility that documentary film material entails (Eriksson 2013, Eriksson P.E. & Eriksson Y. 2015).

Therefore, as in my data, the film editor cannot make use of all the processing possibilities a digital material entails in documentary genres. These conditions limit the film editor’s creative space, regarding to what extent the expressive potential of a digital material, as well as the tool’s capacity can be exploited.

Consequently, contributions to understanding the handling of the challenges with (digital) documentary film material, during the editing process, will free time from problem-solving activities, in favour of creativity in artistic expression. Achieving film continuity, linkage, and rhythm in the flow of documentary sound and image requires time, and dedicated creative efforts that do not intrude on the credibility of the audiovisual material (cf. Reisz & Millar 1968, Orpen 2003, Smith 2005, Pearlman 2009).

Figure 21. The FEAD-modek. Film editing, as an audiovisual design process, consists of framing the design problem, solving it through design decisions, which throughout the work renders audiovisual meaning (potential) for film viewers when they meet the artifact. The film will consist of audiovisual (digital) material that is compiled into an audiovisual object by many sound and image components. This compilation must regard the material’s expressability, its processing, its audiovisual matching, and its semiotics.

Analysis – step 3: care for perceptual aspects

The third step of the analysis, regarding result #3, addresses the use of perception in documentary film editing. During editing documentary film, a significant proportion of time is used to trim an edit point for a satisfactory edit to be made. Perceptual phenomena, such as noise, or visual transients, are often hassles that constrain the continuous audiovisual flow, which can be
time-consuming problems to solve. The more perceptual phenomena at an edit point, the more time consumed, as well as the more re-iterations it takes for the editor to complete the edit to satisfaction. The audiovisual perceptual phenomena considered by the editor relate to what editing principal is used, to the design (shape) of the edit, as well as semiotic properties for each edit.

As shown in step 1 – we can see from the Design Profession Interplay Model (Figure 20) that characteristic activities are of interest when understanding a design profession from a “within”, an academic, as well as a public perspective. Since it consumes significant amounts of time, and affects the expression of the delivered artefact, dealing with audio-visual perceptual phenomena might just stand the trail of being such a characteristic activity for film editors, similar to what previous research has appointed in other areas of design (cf. Smith & Whitfield 2005, Niedderer & Townsend 2014, Adams et al. 2011).

Considering perceptual phenomena, while handling audiovisual material during the editing process, is then also evidently part of the problem-solving, as well as the meaning-making of the film, while creating an expression through the audiovisual matching of sounds and shots at edit points. The problem to solve is how to handle the audiovisual transients, so that the film flows and enhances the desired decoding, while the chosen shots and sounds merge properly, and the viewer’s gaze behaves as intended. These aspects resemble what earlier research has pointed out as delicate in aesthetic problem-solving (cf. Brinck 1999 & 2007, Smith 2005), design meaning-making (Poggenpohl 2009, Ponte & Niemeyer 2013), and audiovisual flow at edit points (Reisz & Millar 1968, Westera 1995, Orpen 2003, Pearlman 2009).

Thus, I claim, that the employment of perception and consideration of perceptual phenomena during documentary film editing is of core interest, since it is both an essential part of the work and important for a satisfactory outcome.

Analysis – step 4: viewers’ perceptual decoding

I will now turn to result four, which regards film viewers’ visual perception of documentary film edits, in the fourth step of the analysis. Viewers watched two versions of the same documentary film sequence: one version made by a film editor to her satisfactory completion (Condition 1); another version in which I had tweaked every edit counter to the editor’s intention with the edit (Condition 2). The viewers’ gazes hit the editor’s appointed AoIs after the edits in both versions, but there was a significant increase in eye movements after the tweaked version (Condition 2), as well as significantly larger pupil sizes.
Figure 22. The FIPA-model of film, attention, intention and perception, connected to the theories of influence of tasks, interpretation of gaze patterns, continuity editing and film continuity. Gibsonian direct perception, perceived continuity, the illusion of continuity in film experience, and continuity support for understanding and memory. The decisive connection point is how perception meets stimuli properties (i.e. the film as an artefact).

When we read these results in light of perception theory, as in the FIPA-model (Figure 22), the film viewing task is the same under both conditions, which cannot arguably render radically different attention. This, in turn, should promote similar gaze patterns for both conditions. Light conditions are also the same for both conditions, and any “second viewing” effect on cognitive load is evened out by the randomized viewing order of the sequences. Thus, differences in pupil dilation can only refer to the presented stimuli. The difference lies in the configuration of the two stimuli, where Condition 1 engenders film continuity, whilst Condition 2 strives for discontinuity in the editing.

What conceivably happens is that in Condition 2, where edits diverge towards discontinuity, attention to the evolving film story is distracted, since gaze directing is broken up. Thus, the support for the viewer’s event detection is disrupted, and possibly also the direct perception of story content. These results confirm what previous research has concluded regarding cognitive load (Kramer 1990, Chen et al. 2016), film discontinuity distracting attention (Cohen et al. 2015, Shimamura et al. 2015), gaze directing in film (Marchant et al. 2009, Cutting et al. 2012), film event detection (Magliano & Zacks 2011), and direct perception at film edits (Gibson 2015 [1979]).

When viewing film, perception is contingent on the film’s time course evolvement, as well as on the visual input. The perceptual disturbances that occur as visual transients in Condition 2 negatively affect the viewer’s time continuity expectations, as well as his/her space continuity expectations,
which together affect the expectations on filmic object’s constant existence, and, thus, inhibit the viewer’s perceived continuity (cf. Smith 2005). Hence, the perceptual reactions of the viewer promote shifts in attention, such as that the illusion of continuity is lost, and the awareness of viewing a film increases.

Furthermore, perceptual reactions that increase cognitive load in film viewers, are known to have negative effects on both understanding the film content, and memory of what was watched. Condition 2 indicates viewer reactions that equal cognitive load at a level of solving difficult problems.


Analysis addressing the second Research Question

Since contextual matters of perceptual precision have now been addressed, it is time to turn to the second, and core research question of this thesis: *In what ways could (audio-)visual intentions be dependent on perceptual precision in the design of continuity edits?*

The aspect of this question that needs further examination is

(5) seizing (audio-)visual intentions, which needs to be understood regarding film editing.

Perceptual precision and documentary film editing as design have been examined in the analysis of the first research question. The relation between (audio-)visual intentions and perceptual precision will be addressed in the *Discussion and Conclusion* chapter below.

Analysis approach and results considered

The analyses regarding the second research question will apply the *FIPA-model*, presented in Figure 22, in the section *Model Application and Interplay*. The interplay of theories concerning the film editor’s intentions and the perception of film edits are from cognitive film theory, and perception studies of moving image edits. The film editor’s intentions regarding viewers’ gazes, where these are stated, are presented in the fifth result.

Analysis – step 5: editorial audiovisual intentions

In this last step of the analysis, a film editor’s intentions regarding the design of film edits, as in result #5, will be addressed. The result consists of two parts:
a mixed method, VOSMET, for capturing intentions regarding the employment of perception during film editing—which will be disregarded now, since it does not, in itself, contribute to answering the second research question; and appointed intentions that the film editor had regarding viewers’ gazes after the edits in the documentary film sequence she made. There are, for most edits, specific areas within the image frame where the film editor intends the viewers to look, and she intends specific audiovisual phenomena to capture viewers’ attention, or not. Most of the time, the film editor is striving for accomplishing continuity edits.

When the film editor has ideas on where the viewer should look across continuity edits, or what should capture the viewer’s attention, such intentions are dependent on a consistent gaze behaviour that focuses the viewer’s attention, and supports a consistent perceived continuity, which creates the illusion of continuity in the mind of the viewer.

As we saw in the fourth step of the analysis, film discontinuity distracts a consistent gaze behaviour away from keeping the viewer’s attention on the story, by inhibiting the perception of continuity, and, instead, increasing the awareness of the editing and creating cognitive load. Therefore, editorial intentions regarding the viewer’s attention, and where to look, require that perceptual disturbances (i.e. audiovisual transients) that distract the viewer, can be kept to a minimum (cf. Smith 2005). Thus, an undisturbed viewing can be reinforced, which is likely to generate a better understanding of the story, as well as a clearer memory of the film content.

Discussion and Conclusion

Any scientific research result, as well as the analysis of the results, must be related to its context of other academic research, in order to make results and analyses comprehensible from a wider perspective. That is the purpose of the discussion, and my results and analyses will now be discussed accordingly.

In this discussion, I will draw on the presented results and consult the theories as presented in the Analysis chapter, in order to answer the research questions of this thesis. I will connect to related research, as well as point towards future research. Finally, I will state the contributions made by this thesis, and draw some conclusions regarding this present research.

In answer to the research questions…

In order to specify the contributions of this thesis, and how they may be understood, its results will now be approached from the perspective of the research questions. Herein, the framing of the research problem of this thesis—whether the film editor’s perceptual precision may lead film viewers’ gazes for stable attention and desired decoding is assessed on its capacity to warrant significant outcomes (cf. Schön 1983). Presented theories will be inserted to shed further light on the results. The discussion will lead to some conclusions, and suggest further research.

From the first result: – Yes, it is appropriate to consider film editing as a design craft that can be further understood as a film and TV craft, if studied from a design research approach. Such perspectives yield theory for a yet deeper exploration of film editing as design work. Public, academic, and from-within-profession perspectives that are applicable to the continued study of film editing, can likely contribute to the understanding of the film editor’s profession (cf. Smith & Whitfield 2005, Adams et al. 2011, Niedderer & Townsend 2014).

From the second result: – Designing, in film and TV production, is dependent on the materiality of audio and video for reasons of creativity. The material comprises perceptual and semiotic aspects that a film editor must take into account. This is part of the film editor’s aesthetic problem-solving, which
is linked to visual perception, audiovisual semiotics, event detection, and continuity/discontinuity in film editing (see Figure 37). The creative processing of the material for a desired expression is therein dependent on the time consumption trade-off between handling a material’s properties, e.g. its audiovisual transients, and its expressive potential. The relation between a digital audiovisual material and its expressive potential, yet under-explored, is a topic that merits further research attention, in order to understand digital possibilities and limitations (cf. Manovich 2001 & 2012).

![Diagram](image)

**Figure 37.** Links between aspects of material handling in film editing (my expanding on Reisz & Millar 1968, Brinck 1999 & 2007, Orpen 2003, Smith & Henderson 2008, Rabiger 2009, Berliner & Cohen 2011, Ponte & Niemeyer 2013), (cf. Figure 16).

**From the third result:** – There are convincing indications that taking perceptual phenomena, such as audiovisual transients, into account is an important aspect of film editing. This result confirms previous research, that considering perceptual factors matters if continuity edits are desired. If continuity is not achieved, it is known to cause perceptual reactions, draw attention to the edit, and increase awareness of the editing, as well as increase cognitive load, which influences memory and understanding of the film (cf. d’Ydewalle & Vanderbeeken 1990, Geiger & Reeves 1993, Smith 2005, Smith & Henderson 2008). How perceptual aspects are part of other (audiovisual) design work is a huge future research topic, with a variety of fields, dwellings, and situations, from which it could possibly be approached, since perception is already taken as a core factor in design (cf. Myers 1989, Simon 1996, Norman 1998, Suwa & Tversky 2002, Santella 2005, Cross 2006, Ware 2008, Koskinen et al. 2011, Maiocchi 2015).
From the fourth result: — Central objects in the film image are not easily missed by viewers due to poor continuity editing. Nonetheless, poor continuity edits make viewers’ gazes more active after edit points, which disturbs attention, and is likely to affect cognition. Pupil sizes also increase significantly, which indicates cognitive load, which, in turn, has an adverse influence on understanding and memory (cf. Kramer 1990, Smith 2005, Seeber 2013, Chen et al. 2016). Other audiovisual properties, such as color, contrast, lighting, rhythm, or ecological distractors could also be further researched with regard to audience response (cf. Pearlman 2009, Süßfüll 2010, Lind-Valdan 2013).

At this point, we can answer the first research question accordingly:

**Attaining a perceptual precision in film editing that frames viewers’ gazes for a stable attention is a design matter of reaching the desired decoding, or not.**

From the fifth result: — The film editor’s intentions regard where on the film image she wants film viewers to look, as well as how and when their attention should be captured. She tries to employ the viewer’s attention both for continuous perception of the story, and for ample perception of scene changes. Intentions regarding perception and attention to (audiovisual) design are still an interesting area for future research, with many possible applications that would render further understanding of design professions and design activities (cf. Lawson 1990, Cross 2006, Koskinen et al. 2011, Aspelund 2014, Niedderer & Townsend 2014, Maiocchi 2015).

From here, we can answer the second research question accordingly:

**When perceptual precision is low, intentions regarding perceived continuity are more likely to fail than when perceptual precision regarding film continuity is high.**

In the case of continuity film editing, editing intentions are linked with viewer decoding of a film story in terms of comprehending its narrative. Such comprehension depends on a non-disrupted viewer gaze across film edits, which the film editor supports by striving to frame the viewer’s gaze to the parts of the moving image that s/he considers contains the most important details. If the viewer’s gaze is kept stable, or framed, at that position across an edit, the decoding is likely to continue undisturbed, with full attention to the story.

Thus, framing the gaze is a matter of accomplishing (audio-)visual design intentions, through considering perceptual phenomena in continuity film editing. How audiovisual designers employ perceptual precision and the role of
perception in audiovisual design must then be understood as core to the realization of design intentions. Perceptual precision is necessary to frame the viewer gaze, which directs attention to intended details.

These answers should encourage future research on audio-/visual perception in yet other design fields where viewer or user perception is considered important. The methods used here may be developed and applied elsewhere, while still capturing the core data, regarding the designer’s intentions, as well as use of perception, along with data that tells what, actually, is being perceived by the target group for the design.

Some implications of these answers

I will now turn to the previously presented research for a discussion of some implications conveyed by the above answers to the research questions.

Design comes with intentions, whereas art might sometimes not. Crafts are understood to inherit properties from both art and design (cf. Niedderer & Townsend 2014). Therefore, regardless of whether film editing is considered a craft, or, as I prefer, an audiovisual design profession, it will be most fruitful to regard it as subject to intentions. These intentions, at times, might draw on expressiveness, or, as is more usual, on narration and comprehension. In any case, film editing will always be carried out through compilations of sound and moving images that carry perceptual properties, as well as semiotic potential (cf. Figure 11). When the semiotic potential, or the perceptual

![Figure 11. Model of film editing, its core working principles (cf. Reisz & Millar 1968, Orpen 2003, Pearlman 2009).](image-url)
properties of the audiovisual material resist the editing intentions of the film editor, s/he is forced into a kind of aesthetic problem-solving in order to withstand the intentions, while considering perception and semiotics.

Aesthetic problem-solving thus involves intellectual considerations in judging semiotic properties and intention outcomes, as well as bodily involvement in regarding perceptual responses. These concerns are addressed by the film editor through a re-iteration process where each edit and its surrounding shots and sounds are assessed by means of active alteration of frames and bits of sound. Hence, aesthetic problem-solving in film editing is activity-based. Thereby, I claim that film editing, as audiovisual design, must be understood by regarding its activities, as other design professions or crafts must be understood. I share this view of design professions with several other researchers (e.g. Schön 1983), and this approach applies to multiple perspectives, such as design professionals’ self-understanding (Adams et al. 2011), public understanding of design professions (Smith & Whitfield 2005), as well as academic decomposition of knowledge in design professions (Niedderer & Townsend 2014).

Understanding design activities includes linking design intentions with thinking and doing during the design process. The film editor’s re-iteration of edits is an activity that aims at adjusting audiovisual transients to a desired flow (smooth or harsh), which means identifying unwanted properties and features of an edit, or bad edit flow (cf. Alexander 1964 p.53-55). Identifying rhythm and other emotions are also important film editing aspects to address. These activities, I believe, represent knowing and reflecting “in-action”, as Schön (1983 p.49-51, 69) phrases it. To enunciate these activities as core efforts by a film editor will add to the topography, or public understanding, of film editing as an audiovisual design profession (cf. Smith & Whitfield 2005), as well as to film editors’ self-understanding (cf. Adams et al. 2011). These activities would be an addition to the activities appointed in earlier film editing literature (Lindgren [1948]2011, Reisz & Millar 1968, Fairservice 2001, Murch 2001). Further exploration of audiovisual design activities, not least in film editing, should therefore be the subject of future research.

Aesthetic problem-solving in film editing
The aesthetic problem to be solved partly depends on what expression is to be achieved. The expression of a film edit resides on a scale between slipping by unnoticed, and manifesting itself in the face of the viewer. Film edits that are made to be noticed, or even reacted or responded to, are considered expressive, whereas smooth edits are imperceptible (Orpen 2003). As film time runs and an edit appears, edit properties can be translated into a matter of a time-graphical event, where the end of one shot meets the beginning of the next, as two different spreads of graphical properties across the screen (Pearlman 2009). When the edit is a cut, this graphical event occurs as the meeting of two frames of distinctly different derivation (except for “jump cuts”). The
time factor allows for an experience of movement of these graphical properties, since moving image frames flow by at least 24 frames per second. If the differences of graphical properties, perceived objects, and/or movement are relatively small, there is linkage between the two shots, and when the differences are large, it constitutes a collision between the shots (Pearlman 2009). The film editor uses linkage and collision as the facets of editing expression.

The commonly used notions of film editing as either continuous or discontinuous (Reisz & Millar 1968, Crittenden 1995, Fairservice 2001, Orpen 2003, Smith 2005) to a substantial extent rely on the linkage–collision scale of the appearance of film edits (see Figure 38). Discontinuity edits frequently use graphical collision as a means of achieving expression, but can also make

![Diagram](image)

**Figure 38.** *Editing Expression Model* of how film editing expressions span over the editing principles axis (continuity – discontinuity), as well as over the graphic appearance axis (linkage – collision). The major tendencies are that edits striving for smooth narration and comprehension are found combining continuity principles with graphical linkage, whereas expressiveness edits combine discontinuity with graphical collision. Examples of exceptions are to be expected.
use of changes that are graphically subtle, yet palpable to the viewer (e.g. “jump cuts”). Continuity edits, on the other hand, frequently employ graphical linkage to make the edits invisible, while also using other means, such as object continuity, continuous movement, and more, across cuts to disguise them from the viewer. The expression of a film edit, passing unnoticed or challenging the viewer, lies, partly, in its employment of linkage or collision, and, partly, in whether it strives towards continuity or discontinuity. The extent of an edit’s expressiveness, then, is dependent on properties of the collision/discontinuity domain. Whereas narration and comprehension depend on linkage and continuity in order to maintain viewer attention and the solid decoding of a story (cf. Figure 38). Either way, the film editor accomplishes the expressions through aesthetic problem-solving activities, as part of pursuing communication.

Multimodality and audiovisual transients

Maintaining attention and uninterrupted decoding of a story is a matter of supporting a communication process aimed at understanding its topic or plot, so that the film viewer is favored in interpretation of its content. Alternatively, if expressiveness is desired to generate some other viewer experience, it is still a matter of communication with an intended expression in mind. Since the content is encapsulated in both sound and image together, the communication taking place must be understood as multimodal (Kress & Van Leeuwen 2001, Ponte & Niemeyer 2013), as must the aesthetic problem-solving when altering sound or image details (cf. Figure 11).

As multimodal in its expression, film editing must adhere to how sound and image affect each other, or rather, how the viewer’s visual and audio perceptions are simultaneous and interact. An audiovisual flow that adheres to film continuity principles, where sound and image are well matched, is known to encapsulate attention (Geiger & Reeves 1993, Cohen et al. 2015), or “keep [it] in control”, as Münsterberg phrases it (1902[1916]), whereas discontinuity distracts attention. This distraction of attention is provoked by the audiovisual transients that come with discontinuity (and graphic collision), since they grab attention away from a continuing audiovisual flow. Such attentional disturbance may affect only the viewer’s focus, e.g. a slight camera “jerk” may appoint a new direction of attention, which is likely to undermine both understanding and memory.

More severely, audiovisual transients might also carry semiotic properties that might invite the viewer to interpretation, e.g. if a small beep noise occurs, it could be taken as a signal of something, which then will be included in the cognitive processing, and the viewer’s meaning-making (cf. Krippendorf 1995). Such audiovisual transients are likely to draw viewer attention away from the audiovisual flow. When film viewer’s eyes are tracked, the reactions
to audiovisual transients at edit points can be captured as compensatory eye movements, as well as larger pupil sizes, after edits.

Whether a film editor intends an edit to be of a continuity or discontinuity kind, it matters significantly if the audiovisual transients are controlled, or not. The film editor achieves such control by means of perceptual precision, controlling the transients, so that they work in accord with the editing intentions. For instance, adopting edits to how visual events evolve on the screen can support continuity by cutting at the end of events, whereas a discontinuity effect comes with a cut in the middle of an event (Magliano & Zacks 2011). In this sense, making events flow, or be disrupted, contributes significantly to what rhythm the film viewer feels (cf. Pearlman 2009, Cutting et al. 2012, Cutting 2014). The film viewer’s gaze follows events, and anticipates the next event with a prepared gaze, when uninterrupted by distractors (Hochberg & Brooks 1996, Smith 2005). If a visual event is interrupted, the viewer’s gaze is more likely to react with compensatory eye movements, and, thus, the attention will be disturbed.

Alternatively, when an audiovisual query draws attention to what will happen next, a cut to the next shot can answer that query (Hochberg & Brooks 1978), or, as in my study, the editor repeatedly uses a query to lead on to a different sub-sequence where other things happen. The viewer’s gaze is likely to follow the lead of the query. How a film editor uses his/her perceptual precision is a matter of what the intentions are, but since perceptual precision is a matter of directing the viewer’s gaze, or not, it needs careful adherence.

Materiality concerns and consequences

A careful adherence to perceptual precision at film edits frequently comes with the need for tryouts of different audiovisual matchings of sounds and image frames. Such iteration is supported by the materiality of sound and moving images of a digital kind, and computer software tools, which allow for countless re-iterations. The digital possibilities of handling audiovisual transients, and other material processing, are many, but rely on user skills and tool capacities, and a proper digital material (Paper B). When these conditions are satisfactorily met, the film editor’s material creative space is still dependent on time for creative processing of the material. Many re-iterations consume time. In my understanding, the available processing time must preferably be used on other aesthetic considerations, such as shape, color, rhythm, and more, rather than on hassles with audiovisual transients, preferably spending maximum effort on the expression itself. Understanding the importance of perceptual precision, and knowing how to identify audiovisual transients, will speed up the film editor’s processing of these transients, so that more time becomes available for dealing with the expression.

At this point, we must acknowledge that audiovisual material is also subject to credibility and viewer preference (cf. Eriksson 2012, Eriksson & Eriksson
A documentary story requires that its material is not processed beyond the degree where its authenticity is maintained. This condition limits expressability, and makes the balancing of how much to process the material more subtle (cf. Figure 9). For instance, too expressive film edits, where cuts are apparent and felt, might be taken as effects, and thus be decoded as a rhetoric approach, potentially undermining credibility.

Figure 9. The material quality of moving images matters to the possibility for designers to process the image and thus express things (Manovich 2001 & 2012, Swenberg 2012). However, material quality also matters for viewers’ preference and ascription of credibility (Eriksson P.E. & Eriksson Y. 2015, Eriksson 2013).

Taken together, these issues of by which means, and how much, to process an audiovisual material are ultimately a matter of sensitivity towards the media and information literacy (MIL) of the viewer. Communication and information are embedded in the audiovisual forms I have addressed, beyond language and text. The viewer’s MIL is not to be taken for granted, but is something that future education must address. How processing of digital materials, as well as perceptual precision, becomes a part of MIL and affects viewer preferences are issues that future research must take seriously. Audiovisual communication must be better understood in order to enable the everyday user of audiovisual communication (cf. Aspers et al. 2004).

Bridging perceptions

A starting point for addressing how audiovisual communication is experienced by the film viewer may be found in Berliner & Cohen (2011), who justify the film viewer’s perceived continuity as replacing the real world in his/her mind. The film, as an audiovisual object of communication, can be provided with audiovisual properties that match each other, either to support a replacement-of-the-real-world experience, or to avoid such an experience of replacement (cf. Grodal 2009). In the replacement experience, the viewer’s perception works in a similar way to when the real world is experienced, which requires perceived continuity, as a kind of direct, action oriented, perception
of the audiovisual world (cf. Gibson 2015[1979]).

Such direct perception must be unconstrained, which again requires perceptual precision in the design of the audio-visual object, in order to avoid audiovisual transients that might stand out, and thus provoke attention to, and awareness of, the film-viewing situation, and provide competing details to attend to. Ultimately, how perceptual precision is a core feature of film editing, as audiovisual design, and consideration of perceptual phenomena one of its core activities, responds to the justification of the filmic (audiovisual) world as replacing the real-world-experience of the viewer.

If this direct-perception approach to film viewing uses a cognitive stance of bottom-up functions in the human perceptual system, film editing also employs perception from a cultural perspective notion, as decoding dependent on intention. First, as we know from the presented results, the film editor has intentions regarding where on the screen the viewer should look. These results align with previous research (Goldstein et al. 2007, Marchant et al. 2009). These editorial intentions succeed the intentions of presenting an audiovisual story that supposedly is comprehensible. Second, these intentions aim at involvement of the viewer’s interest, to the degree that the viewer adopts an intention of his/her own of following the story to the fulfilment of the engendered interest (cf. Grodal 2009). From this perspective, attention and perception in film viewing stems from the film viewer’s intention in continuing watching the film: what meets the curiosity, and nourishes the continuous interest in the film will have precedence to be perceived. This top-down function of perception is thus a part of the viewer’s decoding process.

Since top-down and bottom-up perceptual functions work in parallel as separate cognitive processes (cf. Pinto et al. 2013) that compete for privilege to attention, film editing needs to adjust audiovisual expressions to what is meant to be perceptually top-down processed, and by not being disturbed by bottom-up perceptual inputs (transients). That is, comprehension and narration being supported by continuity editing (cf. Hochberg & Brooks 1996). Likewise, when bottom-up perception is eligible to access attention, top-down processes should be invited to recess. That is what is preferred when expressiveness is strived for through discontinuity editing. Audiovisual transients are employed to stand out as attention-catching effects, making attention shift away from the story.

Awareness of top-down and bottom-up perception in film editing is a matter of control of viewer attention: attention to the story, or attentional flight driven by audiovisual transients. This control of attention requires that the editor is capable of precision mastering of perceptual responses. Thus, the balancing of top-down verses bottom-up attention is a core feature to the role of perception in audiovisual design (cf. Batty et al. 2015). An increased understanding of perception in audiovisual design finds its utmost consequence through the employment of perceptual precision.
I suggest that acknowledging the interplay of top-down and bottom-up perception in relation to attention regarding designed objects, as it is expressed through perceptual precision, bridges a gap between two directions of understanding perception in design: on one hand, addressing humans as cognitive beings, and, on the other hand, seeing them as cultural beings. In design, I believe we must be able to integrate these two perspectives in our understanding of human perception and attention, which, tentatively, should apply to many kinds and fields of design.

**Film editing and perception in design**

Finally, I will turn to how I think human perception and attention relate to the case of studying documentary film editing, from a design research perspective, and make some brief suggestions for future research.

When Rhéa (2008) states that “design both articulates and changes perceptions of the world”, we can see a similar thing going on in film editing: perception is core to how film edits are designed, and the design of the edits is made according to how they will be perceived. And there is convincing evidence that the perceptual precision with which the film edits are designed matters to how they are actually perceived. We can also see that the (audio-) visual thinking that the film editor brings forth through the presented editing intentions consists of concerns for “acts of attention” (cf. Ware 2008) on behalf of the film viewer: s/he is repeatedly enticed to gaze at specific spots of the image frame before and after edit points. The editor urges the viewer to follow a firm lead through the audiovisual stream in order to support comprehension of what is narrated. The viewer’s attention is otherwise easily distracted and comprehension is lost, in a similar manner to what Simon (1996) claims about “focus of attention” as “the key to success” in design.

The reason for this is similar to that of the design context, whereby, as Norman (1998) argues, the human attention capacity is limited. Any small distraction might capture attention and lead it astray. Therefore, instead, the film editor wants to direct the viewer’s attention to what is central to the story, just as a designer needs to direct perception to the core of a design (cf. Smith & Schank Smith 2015).

This is not easy, and Lawson (1990) suggests that the hardest thing about visual representations in design is to render a proper detail perception, where all that is important is well understood. Ware (2008) advocates highlighting of important details. Instead, I suggest that it is the competing details of minor importance that need to be remedied, since these details disturb direct perception of the thing to be perceived (cf. Gibson, 2015[1979]). Such care for disturbances is similar to what the film editor does as s/he removes all that disturbs attention, to keep it “in control” (Münsterberg 2002[1916]). Perhaps, too, a better understanding and application of perceptual precision can
offer a solution to detail perception in other areas than audiovisual design. I suggest that this can be done as a possible direction for future design research (cf. Koskinen et al. 2011). Furthermore, perceptual skills might thus become part of the topography of other design professions as well (cf. Smith & Whiftield 2005).

Contributions

This thesis is a contribution to research on (audiovisual) design work and perception in design. Evidently, its primary contribution regards the understanding of film editing, how perception is employed with precision during editing, and to what end.

In the wider, general, academic perspective, this thesis contributes to the continuing investigation of professions, and it attempts at research across academic fields and disciplines, such as media and design. Further, the thesis contributes to the ongoing development of communication research, where audiovisual communication draws greater interest as these forms of interpersonal, as well as mass communication continuously extend by the introduction of new artifacts and applications. The particular contribution of the thesis is its reasoning about the formation of audiovisual knowledge, as this is a prerequisite for such communication, and how audiovisual knowledge is applied in the professional communication context.

Another contribution of the thesis is to the understanding of audiovisual knowledge from an epistemological perspective. The way perception is applied during (audio-) visual thinking adds to the understanding of sensory integration of spatial and temporal inputs into a continuous flow of impressions that a human uses to understand the current condition of the surrounding world. A world that may be substituted by audiovisual inputs in the form of, for example, film.

More specifically, this thesis contributes to the understanding of Film and TV production from a new perspective, namely, design. A further claim is that the design perspective is most fruitful in trying to understand what actually goes on when sounds and moving images are being created, integrated, and produced during industrial film and TV production processes. Thereby, my thesis endeavors to assimilate film and TV production research into design research.

Most specifically, this thesis contributes to research on innovation and design by its mix of methods, and a supporting methodology, that, in tandem, addresses the creation of new artifacts, the thinking going on in such a process, as well as the human perceptual response to the created artifact. This approach could inspire future colleagues in the field in researching creation processes.
Conclusion

There is hitherto design and craft knowledge that prevails as un-articulated, which this thesis attempts to rectify, by way of articulating perceptual skills as core to audiovisual design in the form of film editing. I have met the objective to link film editing intentions with the film editor’s efforts to frame the viewer’s gaze to important visual details, and point out how such framing is important in the understanding of a film. In doing so, I have targeted the role of perception as an inherent aspect of audiovisual design, as well as how film editors employ perceptual precision as a means of realizing their intentions when designing. Thereby, I have demonstrated how perceptual skills can reside integrated in audiovisual thinking and audiovisual knowledge.

My conclusions are, firstly, that the current framing of the problem – how the film editor’s perceptual precision can lead the gaze of the film viewer for stable attention, and desired decoding – does lead to significant outcomes:

(I) understanding how visual perception is employed;

(II) understanding different ways which visual perceptual phenomena are dealt with in film editing;

(III) understanding how perception and perceptual phenomena are inherent aspects of the process of designing film edits.

Thus, the framing of the research problem used in this thesis stands the problem-framing test (cf. Schön 1983).

In this thesis, I have monitored how film editing, as an audiovisual design profession, employs perception as core to its (audio-)visual thinking in achieving its design results. Perceptual precision is a means of realizing audiovisual design intentions in terms of expression and function, by linking these intentions with the film editor’s efforts to frame the viewer’s gaze to what is estimated to be most important. In contrast to Ware (2008), who emphasizes the steering of perception by highlighting important (graphical) features, I recognize the diminishing of competing (audiovisual) features as the decisive aspect of framing the viewer’s gaze when creating moving images. Framing the gaze is a matter of balancing top-down against bottom-up perceptual processes that a film viewer is involved in, and the film editor has to identify bad-flow edits with harsh audiovisual transients and re-iterate these edits until they appear smooth to the eye.

The importance of perceptual precision as core in film editing highlights perceptual skills as a form of knowledge, regarding both audiovisual matters, and design. Further, perceptual precision is shown to matter to how viewers decode what they see and hear while watching a film. As such, perceptual precision becomes a core function of (audio-)visual thinking in design.

Audiovisual designers are perception directors, which should therefore influence their self-understanding, as well as academic and public understanding
of such professions. The current outcomes, I claim, should affect audiovisual (and perhaps other) design curricula where (audio-)visual thinking is employed (cf. Pratt 2011, Costa et al. 2014), and also support inexperienced film editors, as well as other audiovisual designers in developing their set of audiovisual design skills.

An increased understanding of what (kind of) knowledge film editing is founded upon is a necessary departure for bringing the profession out of its tradition of silence (cf. Murch 2006). Articulating knowledge that has hitherto been un-articulated, is an empowerment for film editors. This is audiovisual knowledge that occurs “in-action” (cf. Schön 1983), when the film editor assesses the audiovisual flow of shots and sounds during the becoming of a film.

The key audiovisual knowledge by means of which film editors can master the audiovisual flow of a film, is the perceptual precision they use to control bottom-up perception of audiovisual transients, in relation to viewers’ top-down perceptual interest in the film story, and thereby assess which edit flows, and which does not, according to intention.

I rest with the conclusion that the cut where the pint-of-beer shots met, had very few and small visual transients, and none in audio, that could possibly trigger any bottom-up perception competing with top-down perception of following the action of the bar tender pouring and serving that drink. Rather, the movement going on in the shots were major visual transients that engaged bottom-up perception in the same direction, as a complete and strong event that the gaze was framed to follow, supporting the ongoing top-down perceptual processes, to the degree that the edit was completely disguised. The two shots became one continuous audiovisual flow in my own perception, and there was nothing perceivable to separate them.
References

Adams, Robin, Daly, Shanna, Mann, Llewellyn, & Dall’Alba, Gloria (2011). Being a professional: Three lenses into design thinking, acting, and being. Design Studies, 32(6), 588-607.


Koskinen, Ilpo, Zimmerman, John, Binder, Thomas, Redström, Johan, & Wensveen, Stephan (2011). Design research through practice: From the lab, field, and showroom. Waltham, MA: Morgan Kaufmann.


# Appendix 1

## Tables

### Table I. Perceptual phenomena, with codes and defining criteria (Paper F).

<table>
<thead>
<tr>
<th>Phenomenon:</th>
<th>Code:</th>
<th>Criteria:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some image feature or object is an explicit candidate intended to capture attention</td>
<td>Attention Capture Intention</td>
<td>The editor states an intention for an image feature or object to capture attention (or overtly makes it salient)</td>
</tr>
<tr>
<td>Some aspect of the film draws attention to that something is about to come (e.g. occur)</td>
<td>Visual Query posed</td>
<td>A sound or image feature indicates an event, and thus provokes curiosity on an attentional (basic) level</td>
</tr>
<tr>
<td>Some aspect of the film meets attentional expectations of what was to come</td>
<td>Visual Query Answered</td>
<td>A sound or image feature corresponds to the curiosity provoked by the previous image</td>
</tr>
<tr>
<td>Some aspect of the film exploits attentional expectations of what was to come*</td>
<td>Visual Query Exploited*</td>
<td>A sound or image feature makes use of the curiosity provoked by the previous image to capture attention*</td>
</tr>
<tr>
<td>A specific object or area within the image frame is the prime candidate to capture attention due to its saliency, after an edit</td>
<td>Point of Interest</td>
<td>There is an overtly salient object or area within the image frame that is the prime candidate to capture attention (or stated by the editor as such), after an edit</td>
</tr>
<tr>
<td>Some aspect of the film draws (pulls) or pushes perception (often in a direction) before an edit</td>
<td>Cue Type</td>
<td>There is an overtly salient sound or image feature that either pulls the eye to it, or pushes the eye away (to the middle of the screen), or triggers attention</td>
</tr>
<tr>
<td>Some part of the image, that seems not to change, draws attention to itself across an edit point</td>
<td>X’ aimed at Inattentional Blindness</td>
<td>There is an overtly salient part of the image that seemingly does not change across an edit point, and that draws (focused) attention to itself</td>
</tr>
<tr>
<td>New information presented in the film draws attention just before an edit point</td>
<td>X’ aimed at Attentional Blink</td>
<td>Viewer’s internal attention is directed towards just extracted information (maximum 500ms before edit point)</td>
</tr>
<tr>
<td>A visual event ends or starts just before an edit point, which draws attention to itself</td>
<td>Visual Event Segmentation</td>
<td>The segments of a visual event are employed as edit points, in order to exploit viewer’s attention (to events)</td>
</tr>
<tr>
<td>Spatial continuity is upheld across an edit</td>
<td>3D-Continuity Match</td>
<td>The impression of a continuous 3D space is withheld across an edit</td>
</tr>
<tr>
<td>Condition</td>
<td>Technique</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Shots have a similar graphical layout with their respectively most salient or interesting spot on the same part of the image surface, across an edit</td>
<td>Graphical Match</td>
<td>Shots are matched across an edit in order to maintain the viewer’s gaze on the same spot of the image surface</td>
</tr>
<tr>
<td>There are no (or only small) visual transients across an edit</td>
<td>Visual Transient Avoidance</td>
<td>An edit is shaped to avoid or minimize possible visual transients</td>
</tr>
<tr>
<td>Time continuity is upheld across an edit point</td>
<td>Time Continuity Match</td>
<td>The impression of time passing continuously is withheld across an edit point (uninterrupted events)</td>
</tr>
<tr>
<td>Movement of the focal object, just before an edit point</td>
<td>Focal Object Movement</td>
<td>The edit point is chosen so that the focal object moves just before, and thus starts an event that the viewer attends to, and expects to unfold uninterruptedly</td>
</tr>
<tr>
<td>Saccade or blink provocation occurs just before an edit point</td>
<td>Saccadic Suppression or Blink</td>
<td>The edit point is chosen so that a visual event that provokes a saccade or a blink occurs just before</td>
</tr>
<tr>
<td>Saliency of focal object at a non-cued edit</td>
<td>Non-cued: Focal Object Salient</td>
<td>Across a non-cued edit there is a salient focal object in each shot which steers viewer’s time expectations</td>
</tr>
<tr>
<td>Relocation of focal object at a non-cued edit</td>
<td>Non-cued: Focal Object Relocation</td>
<td>Across a non-cued edit a focal object is relocated between shots, through different graphical layouts, which steers viewer’s time expectations</td>
</tr>
<tr>
<td>Matching of the focal object’s relocation to the viewer’s attentional set, at a non-cued edit</td>
<td>Focal Object Reloc. matches Viewers’ Attentional Set</td>
<td>The relocation of focal object matches viewer’s attentional set, at a non-cued edit</td>
</tr>
<tr>
<td>Shift of background at an edit point</td>
<td>Background Change</td>
<td>The background of the image changes at the edit point</td>
</tr>
</tbody>
</table>