Tampere University

Mika Kuitunen

DESIGNING CONTROL SURFACES FOR VIDEO EDITING

A Human-Centred Approach

Master of Science Thesis
Faculty of Information Technology and Communication Sciences
Examiners: Dr. Kirsikka Kaipainen
Prof. Kaisa Väänänen
23 April 2022

ABSTRACT

Mika Kuitunen: Designing Control Surfaces for Video Editing
Master of Science Thesis
Tampere University
Computing Sciences - Human-Technology Interaction
23 April 2022

Video editing is the task of compiling media of various types to form a cohesive piece that tells a story. While simple video editing can be done even on modern smartphones, complex projects containing media from multiple sources and requiring more processing power are done on desktop with non-linear editors (NLEs), editing software that allows the user to easily adjust any part of the sequence. This is a task that requires both creativity and technical knowledge to solve the problem of storytelling and convert the idea into the final cuts. A control surface is a physical user interface that facilitates the use of a specific software, and for video editing the most common control surfaces are the keyboard and mouse.

This Master's thesis examines video editing control surfaces from a human-centred perspective through a human-centred design process. The research problem consists of three research questions: (1) What is important to the user in a video editing control surface, (2) What is the right degree of specificity for a control surface and (3) What design considerations does the task of video editing cause. The questions were answered by developing a new control surface concept using the human-centred design process, and evaluating both the process and its results.

Initial context were gathered via a literature review and a contextual inquiry, which produced personas for the potential user and some tentative heuristics that could be used to evaluate the UX of existing control surfaces. The heuristics were developed into a full set through a formalized process and three different control surfaces were evaluated to find potential issues and ideas for a new product concept. A prototype of a new control surface was created and evaluated through an online survey, and finally, the design guidelines for video editing control surfaces were formed.

The results of the research indicate that there might not be a need for editing-specific control surfaces as the keyboard and mouse are sufficient for most users, even experienced professional editors. The actions of modern NLEs are designed to be completed without specialized control surfaces and for many tasks, a button press is already an efficient interaction. The users are willing to spend time building muscle memory for their tasks, so the creative problem-solving becomes the bottleneck instead of the execution of each action.

The heuristics created during the research worked well and could be used to evaluate physical user interfaces in other contexts as well. If a new video editing control surface is designed, the most important considerations are customizability, well-working combination of hardware, software and integration with the NLE, and allowing the user to memorize the most common actions while being able to find less common actions easily. The users' deep level of expertise with their current software affects the results and while novel interaction types were not found to be interesting in this research, the result may change if NLEs with better support for new controls are created.

Keywords: video editing, control surface, controller, physical user interface, human-centered design, heuristic evaluation, design guidelines

TIIVISTELMÄ

Mika Kuitunen: Videoeditoinnin käyttöpintojen suunnittelu
Diplomityö
Tampereen yliopisto
Tietotekniikka - Ihmisen ja teknologian vuorovaikutus
23. huhtikuuta 2022

Videoeditointi on tehtävä, jossa erilaista mediaa videosta grafiikkaan yhdistetään kokonaisuudeksi, joka kertoo tarinan. Yksinkertaista videoeditointia voi tehdä jopa älypuhelimella, mutta monimutkaiset projektit, jotka hyödyntävät laajempaa materiaalivalikoimaa ja vaativat enemmän suorituskykyä, tehdään tietokoneella epälineaarisella editorilla. Epälineaarinen editori on ohjelmisto, joka mahdollistaa videosekvenssin muokkaamisen mistä tahansa kohtaa. Videoeditointi vaatii sekä luovuutta että teknistä osaamista, jotta tarinankerronnallinen ratkaisu saadaan muutettua leikkauksiksi. Videoeditoinnin tapauksessa tyypillisimmät käyttöpinnat, eli fyysiset käyttöliittymät tietyn ohjelmiston käyttöön, ovat näppäimistö ja hiiri.

Tämä diplomityö tarkastelee videoeditoinnin käyttöpintoja ihmiskeskeisen suunnitteluprosessin kautta. Tutkimusongelma koostuu kolmesta tutkimuskysymyksestä: (1) mitkä asiat ovat käyttäjälle tärkeitä videoeditoinnin käyttöpinnassa, (2) mikä on oikea erikoistumisen taso käyttöpinnalle, ja (3) mitä huomioita videoeditoinnin tehtävä nostaa esiin käyttöopintoja suunniteltaessa. Tutkimuskysymyksiin vastattiin kehittämällä uusi konsepti käyttöpinnalle ihmiskeskeisen suunnitteluprosessin avulla, ja sekä prosessia että tuloksia arvioimalla.

Tutkimuksen konteksti selvitettiin kirjallisuuskatsauksella sekä kontekstuaalisella tiedustelulla, joista tuotteena saatiin suunnitteluprosessia hyödyttävät persoonat sekä alustavat heuristiikat käyttöpintojen arviointiin. Heuristiikat kehitettiin valmiiksi formaalia prosessia käyttämällä, ja kolmelle erilaiselle käyttöpinnalle tehtiin heuristinen arviointi. Uudesta käyttöpinnasta tehtiin prototyyppi, jota arvioitiin verkossa tehdyn kyselytutkimuksen avulla. Lopuksi luotiin suunnitteluohjeisto videoeditoinnin käyttöpinnoille.

Tutkimuksen tuloksien perusteella erikoistuneille käyttöpinnoille ei välttämättä ole tarvetta, vaan näppäimistö ja hiiri ovat riittävän tehokkaat käyttöpinnat jopa kokeneille ammattilaisille. Modernien editointiohjelmistojen toiminnot ovat suunniteltu tehtäväksi näppäimistöllä ja hiirellä, ja napin painallus on monessa tapauksessa tehokas interaktiotapa. Käyttäjät ovat valmiita panostamaan aikaa lihasmuistin rakentamiseen, jolloin luova ongelmanratkaisu muodostuu pääasialliseksi pullonkaulaksi teknisen suorittamisen sijaan.

Tutkimuksessa luodut heuristiikat toimivat hyvin, ja ovat hyödynnettävissä myös muiden fyysisten käyttöliittymien arviointiin. Uutta videoeditoinnin käyttöpintaa suunniteltaessa tärkeimmät asiat ovat muokattavuus sekä laitteiston, ohjelmiston ja editointiohjelman integraation yhteistoimivuus. Lisäksi käyttäjän on pystyttävä oppimaan yleisimmät toiminnot ulkomuistiin, ja vähemmän yleiset toiminnot tulevat olla helposti löydettävissä. Käyttäjien syvä kokemus nykyisten ohjelmistojen kanssa vaikuttaa tuloksiin, ja vaikka uudet interaktiotavat eivät vaikuttaneet tässä tutkimuksessa kiinnostavilta, tulos voi muuttua, mikäli näitä interaktiotapoja paremmin tukevia editointiohjelmistoja kehitetään tulevaisuudessa.

Avainsanat: videoeditointi, videon leikkaus, käyttöpinta, fyysinen käyttöliittymä, ihmiskeskeinen suunnittelu, heuristinen arviointi, suunnitteluohjeisto

PREFACE

> Remember kids, the only difference between screwing around and science is writing it down.
>
> – *Adam Savage & Alex Jason*
> MythBusters, 13 May 2012

This thesis started out a few years ago as screwing around and started its transformation towards science in the fall of 2021. Thank you to Dr. Kirsikka Kaipainen who acted as my instructor and examiner at Tampere University and provided invaluable feedback and advice during the process.

My time at the university has gifted me both strong and positive experiences, and I am grateful for all the people I have got to meet and the experiences I have got to share with them. Thank you to my friends from NääsPeksi, Rakkauden Wappuradio, TT-kamerat and TEA-club for a memorable journey. Thank you to my family for the support, Sisyphean Films for keeping the spark of creativity alive and the Grande crew for being you.

Tampere, 23 April 2022

Mika Kuitunen

CONTENTS

LIST OF TERMS AND ABBREVIATIONS

Keyboard shortcut	Keys or key combinations which invoke an action without any further interaction, for example pointer movement or menu navigation (*ISO 9241-171:2008*)
Macro	(Computer science) A single instruction that expands into a sequence of actions that the computer software carries out (Oxford Learner's Dictionary 2022)
Metadata	Data providing information about the primary data and describing it, but not the content of the data (Riley 2017)
Modality	(Human-computer interaction) The channels of sensory input and output used by the user interface, most commonly visual, auditory and tactile/haptic (*ISO 9241-112:2017*)
Scene	A dramatic event that usually happens in a single location and time (Dirk 2022)
Shot	The atomic component of a video, a continuous recording of a certain action from a certain perspective (Dirk 2022)
Take	A single recording of a particular shot, often repeated with slight adjustments until the director approves the shot (Dirk 2022)

GUI	Graphical User Interface
HCD	Human-Centred Design
NLE	Non-Linear Editor, software or system designed for the task of non-linear editing
UI	User Interface
UX	User Experience

1. INTRODUCTION

Video editing is both a creative and a technical challenge where the task of the editor is to tell a story with the raw material they are given. By piecing it together in different ways, altering the appearance and sound of the material, and adding their own style to the video, the editor has the power to paint the story in whatever light they wish and even change it completely. An often circulated quote in the film industry, possibly by director Robert Bresson, states that "A film is born three times. First in the writing of the script, once again in the shooting, and finally in the editing" (Hiebert et al. 1968). The power of editing is not limited to film, however: most videos you see are edited in some way, from simple trimming and stitching of social media videos to vast amounts of raw material condensed into a single, high-paced music video.

The topic of this research is complex video editing, meaning work that includes large amounts of raw material and processing requirements, as opposed to editing a short video recorded on a smartphone. This type of work is usually reserved for desktop software and requires some learning in order to become proficient. Video editing and its tasks are defined in section 2.2. More specifically the focus of the research is the physical user interfaces, or control surfaces, used for video editing. These days most editors use a keyboard and mouse for their work but more specialized options exist and many interface types are still unexplored for this domain. The goal of this research is to find out whether there is potential in specialized control surfaces and what are the important considerations of video editing from the user perspective. The primary product of the research is a set of design guidelines for video editing control surfaces. These guidelines are meant to be used during the design process, but an additional set of heuristics are formed for evaluating the usability of a control surface. The research is based on three research questions:

1. **What is important to the user in a video editing control surface?** This is likely to vary between users, their responsibilities in a given project, and the tasks they are currently doing. It is important to define these priorities at a high enough level so that they apply regardless of the specific tasks.
2. **What is the right degree of specificity for a control surface?** The user interface can become too specialized to the point where it is only useful in a single task, but too generic a design might reduce the efficiency of the user interface.

3. **What design considerations does the task of video editing cause?** Video editing is a specialized part of human-computer interaction but also a creative process, and the considerations are likely to change with the experience and the specific tasks of the target users.

There is no formal definition for a video editing control surface, but one can be created by combining the related definitions and the requirements set for the control surface by the research scope. ISO 9241-110 defines user interface as a "set of all the components of an interactive system that provide information and controls for the user to accomplish specific tasks with the interactive system" (*ISO 9241-110:2020*). An interactive system is the "combination of hardware and/or software and/or services and/or people that users interact with in order to achieve specific goals" (*ISO 9241-11:2018*).

The specific goal in the scope of this research is to create a cohesive video from the raw material and the specific tasks to achieve this goal are the tasks a video editor would do in the editing phase. The interactive system in this case consists of the computer, the editing software, and the physical user interfaces used to control the software and receive feedback from it. However, the scope limits the user interface to the physical user interfaces used to control the software and the information and controls they provide directly. The physical user interface does require some software of its own to operate, such as firmware on the device or a companion software that acts as a bridge between the user interface and the editing software. Based on this we can define the video editing control surface as a physical user interface used for the task of video editing, containing the hardware and the software that enables the interface to function with the editing software.

The research is structured as follows: chapter 2 covers the relevant theory, concepts and literature of the topic and the overall research process is overviewed in chapter 3. Chapter 4 covers the contextual inquiry, its results and its effects on the rest of the research. In chapter 5 a set of heuristics is formed for evaluating video editing control surfaces and a selection of existing products are evaluated. Chapter 6 covers the design of a new control surface and the user evaluation of the prototype, chapter 7 compiles the findings of the entire process and the design guidelines are formed, and chapter 8 concludes the research.

2. THEORY AND BACKGROUND

This chapter covers the history of video editing, its evolution to the current form, and describes the editing process. The editing process is discussed from a user experience (UX) perspective, and standards and previous work relevant to this research are explored.

2.1 A Brief History of Editing in Film and Video

At its core, editing is about stringing together a series of separate "shots", recordings of specifics actions, to form a cohesive piece. In the book The Technique of Film and Video Editing (Dancyger 2013, pp. 3–5), author Ken Dancyger states that the concept of editing was invented soon after the first motion pictures were displayed to the public. Since the turn of the 20th century editing has evolved into a complex task requiring both technical expertise and creativity from the editor. The core of editing has stayed the same: assembling separate shots into a sequence, but the tools have changed and expanded drastically.

The medium has dictated the working methods throughout the progress, starting with editors splicing together strips of film manually to form the final cuts (Morris 1999). Editing stations were eventually introduced and the advent of videotape changed the way the recordings were stored, but the important distinction is that these are still linear editing methods: the editor has to start from the beginning of the sequence and work their way through the cuts, appending a new shot to the end of the previous one. This means that changes to the middle of the sequence, such as inserting new shots, changing the duration of a shot, or replacing existing shots are very difficult, and the editor has to visualize the full sequence mentally before committing to the cuts. This is analogous to shooting an entire video with multiple shots on a continuous videotape: you would start by planning your shots and what actions are included in each of them, then proceed to record your shots in the order they should appear in the final video.

In 1970 computer technology was becoming powerful enough to handle video editing and the CMX 600 editor started the movement towards digital media management (Rubin and Diamond 2000, pp. 44–46). Suddenly the editing paradigm shifted from linear to non-linear. Instead of a "master tape" onto which the sequence was recorded, each shot is accessed as needed to play back the content between the "in" and "out" points and the sequence simply keeps track of the order and in/out points of the source material (Rubin and Diamond 2000, pp. 4–5). The editor no longer has to commit to their edits and instead can at any point go back and change the already laid out sequence.

In the next few decades computer-based editing opened up the possibility for editors to alter the picture in many ways and work with sound more than was previously possible (Dancyger 2013, pp. 391–392). One of the first and most successful computer-based editing programs was Avid Media Composer (Avid Technology 1989), first released in 1989 and remaining one of the most popular Non-Linear Editors (NLEs) for professional editors today. In the 1990s the editors were still proprietary systems with customized hardware but over time the direction has been more and more towards computer software you can use on almost any computer with a keyboard and mouse.

This shift from linear to non-linear editing and the adoption of digital post-production tools has brought more possibilities than ever to the hands of the editor, and depending on the project they could consolidate tasks traditionally done by a team of artists to a single role. Over time the physical user interface has become more generic and although editing-specific control surfaces do exist, they are no longer a requirement for video editing.

2.2 Video Editing as a Process

A short explanation of the video editing process and workflow is necessary to understand the user experience aspect of it. The exact workflow and order of operations varies from person to person but the primary steps to get from raw material to a finished edited video remain the same.

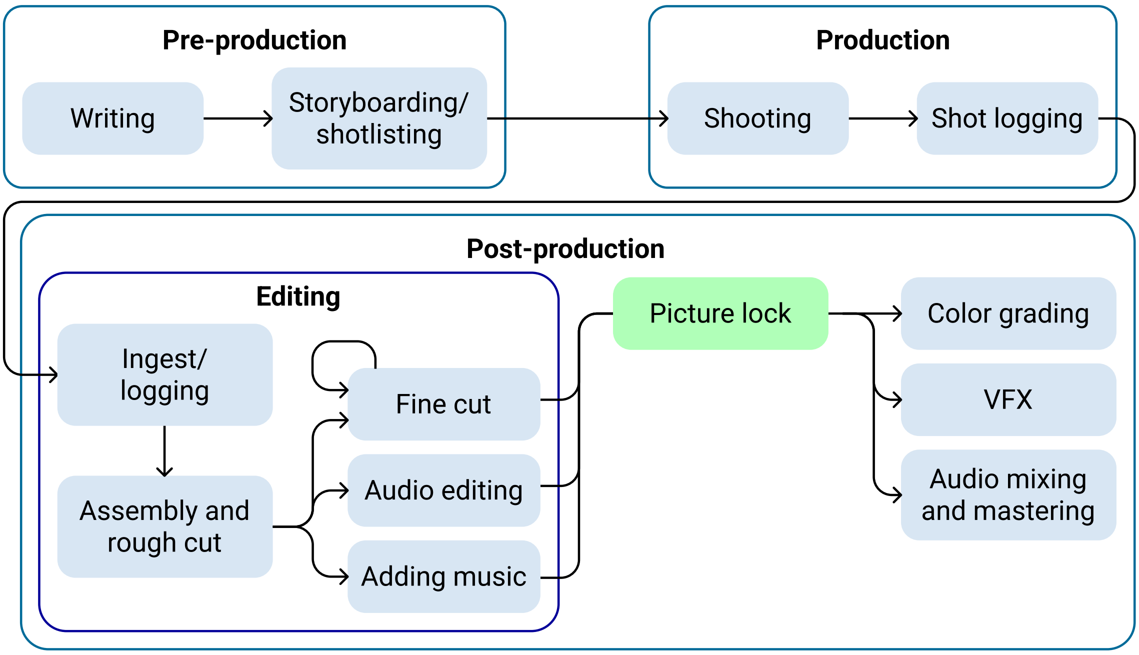


Figure 2.1. *Video production process*

Figure 2.1 depicts the general steps of a video project from pre-production planning to the finished product with editing as the focus. The editing process starts by importing the raw material, which can be video footage, audio, still images or graphics. This phase is also known as ingesting and depending on the size of the project, might include logging the material as well. During logging the editor, or in some cases the editor's assistant, will go through the raw material and log important metadata about the material, such as shoot date, location, scene (of the script), shot and take.

After the ingest phase the editor will assemble a rough cut of the sequence, selecting the best takes of each shot and deciding how to use the shots to tell the story. This phase is often focused on picture editing but includes audio editing as well. If the project has a separate director, they are usually involved in the post-production and discuss changes with the editor. The rough cut is then iterated on until the editor and possibly the director are happy with the edit. The picture is then "locked", meaning no further picture editing will be done to the sequence. At this point multiple steps can be done in any order and by different people: the audio editing is completed and the audio is processed and mixed, the shots are colour graded to improve their appearance and match between separate shots, and any visual effects that need to be done on the shots are completed.

The primary focus of this research is the assembly and editing phase of the process, from having the raw material ready for editing to completing the picture lock including audio editing.

2.3 Video Editing From a User Experience Perspective

Modern video editing is done mostly on computers with generic input devices, the keyboard and mouse. Sometimes application-specific control surfaces are used, but most often these are used in addition to keyboard and mouse, not replacing them completely. Some editors are available on tablets and smartphones where the touch screen is the primary mode of interaction, but this research focuses on the computer-based variants.

The interaction is based on the editing software's graphical user interface (GUI) with visual and auditive output. Usually, the auditive output is reserved for listening to the audio of the material as the user is working on it, and only in rare cases errors are output via audio. The GUI is used more broadly to display not just the edited sequence but the software's state and the tools for manipulating the sequence.

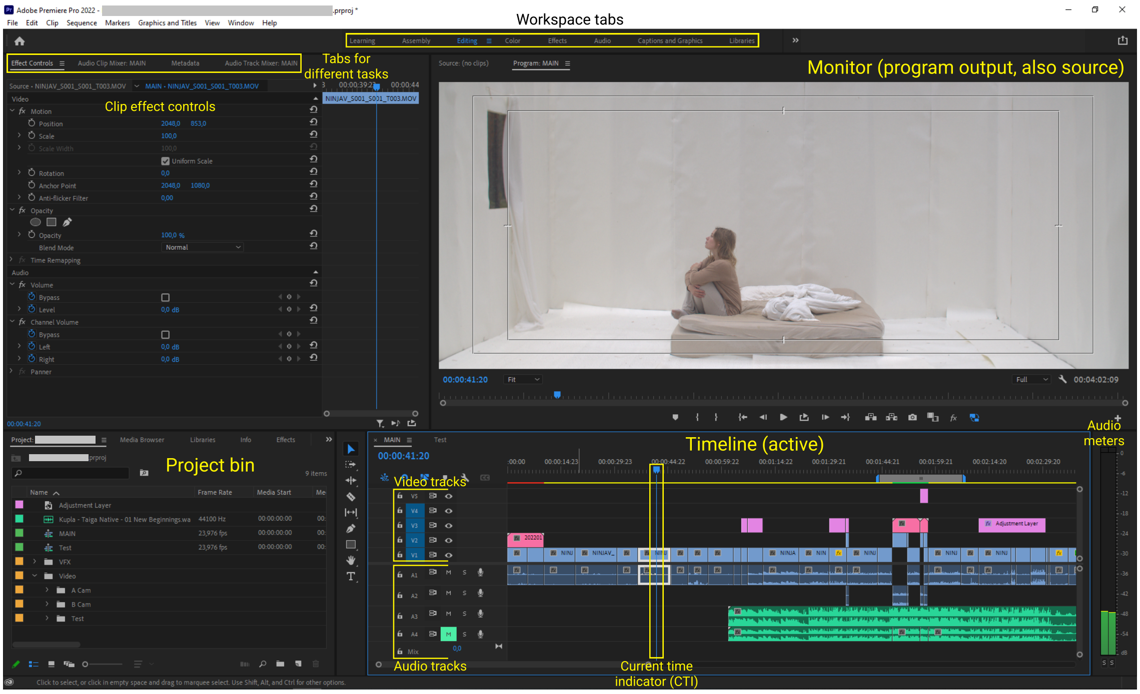


Figure 2.2. *Adobe Premiere Pro CC 2022 user interface*

A screenshot of the current Adobe Premiere Pro user interface (UI) is presented in figure 2.2 and it is representative of other editing software as well. The UI consists of multiple panels that serve different purposes, and they can often be rearranged and resized to fit the users' preferences. The most important panels for the editing phase are the project bin (sometimes called the source bin), the timeline and the output monitors.

Shots can be dragged onto the timeline as-is and edited from there, or viewed in the source monitor (similar to program output) where the editor can mark in- and out-points and insert specific sections of the shot onto the timeline. Once the material is on the timeline, the editor will work mostly with the timeline panel and the program output for the cutting operations, occasionally going to other panels to work with video and audio effects. The editor will do all their operations, such as splitting, trimming and reorganizing on the timeline, while watching the result on the program monitor and listening to the audio of the edited sequence. They can also use tools to analyse both the colours of the picture and the audio levels to ensure they are as the editor intended, such as the audio meters next to the timeline.

Input is done via mouse and keyboard (pointing device and text input) or some more specific hardware that better matches the input interaction to the effect and feedback of the software (e.g. knobs, sliders and tracking balls). The input devices offer limited tactile feedback about the position of the users' hands on the device and combined with the muscle memory of an experienced editor, this allows the user to keep their eyes on the GUI most of the time.

A useful model for understanding user experience is Marc Hassenzahl's UX model (Hassenzahl 2004), which covers both the designer's and the user's perspective. The model is depicted in figure 2.3 and at the centre is the product character, a high level description of the product's attributes. The product's attributes are split into two major categories: pragmatic and hedonic attributes. The pragmatic attributes include manipulation, meaning manipulation of the environment. The functionality for this must be both present and accessible, and so typical examples of this category are useful and clear. The pragmatic attributes help the user in fulfilling their goal. Hedonic attributes include stimulation, identification and evocation. These attributes do not help the user achieve their goals directly, but rather focus on the psychological reaction and the pleasure generated by the use of the product. Stimulation may refer to for example features not yet used but present in the product, thus providing the opportunity for self-development. Identification means the user's self-expression through owning and using the product. Evocation is the act of bringing up memories like past events or thoughts that are important to the user.

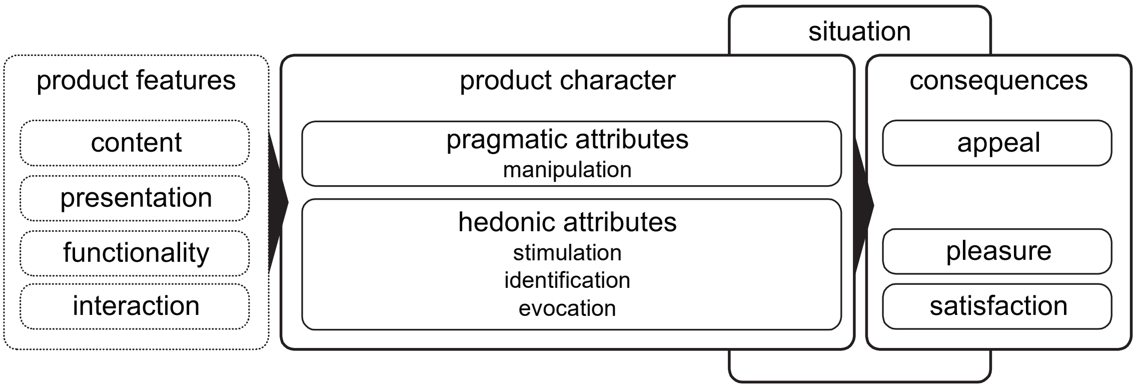


Figure 2.3. *Hassenzahl's User Experience model (Hassenzahl 2004, p. 32)*

The designer selects the features to be included in the product to convey the *intended product character*. The user constructs the *apparent product character* based on the features they perceive, and this might differ significantly from the designer's intent. The apparent product character leads to the consequences: the user's overall experience of the product. As the topic of the research is a control surface for expert use, the focus will be on the pragmatic attributes of the product. The primary goal for the product design is to be useful and usable, but hedonic attributes should not be overlooked even in professional products.

2.4 Previous Research on Video Editing Control Surfaces

Very little research has been done on the topic of video editing from a user perspective, especially on the interaction techniques used to achieve the tasks. A lot of research exists on the subjects of automatically detecting video content and autonomous editing based on scripts or storyboards but as Dancyger states, editing at its core is a creative process and technology can only go so far in helping the editor achieve their storytelling goals (Dancyger 2013, p. 392). Research related to user interfaces and interaction in video editing is covered in the rest of this chapter.

Zigelbaum et al. explored tangible video editing in their research by creating an active token for each shot, meaning the physical token contained a display and the ability to play the contained shot and communicate with other tokens attached to it (Zigelbaum et al. 2007). Their research focused on the assembly and reorganizing of shots into a sequence and the collaborative aspect of video editing. The tangible video editor did increase the active collaboration of the participants compared to graphical user interface-based editing, where one of the participants would often take responsibility of operating the software. The research team also discovered that representing the shots with tokens facilitated the organization of the shots in the workspace, namely the participants would move the shots they considered important close to their bodies. This is an interaction not usually afforded by modern editing software: shots can be organized into bins and labelled with colours, but the user and the shots are not present in the same space. One of the drawbacks of the tangible video editor was the scope as trimming or cutting the videos was not possible. Another important consideration is the lack of a singular "undo" action common in software. If the participants tried something they didn't like, returning the sequence to the previous state would take multiple steps and the history of actions would have to be in the users' memories.

Merz et al. expand on the research done by Zigelbaum et al. in their research (Merz et al. 2018). Their prototype, ClipWorks, is a tangible user interface for video editing like the Tangible Video Editor. Where ClipWorks differs is instead of active tokens, the system uses passive tokens placed on an interactive surface that reads the position of the tokens and projects a digital user interface onto the surface. Editing operations are done by adding, removing, reorganizing and moving tokens and different tokens correspond to different entities or actions, such as a video token and a play token. ClipWorks also adds the ability for the user to import their own video material for editing. This design is primarily aimed towards late elementary school or early middle school students, and it was made to facilitate collaboration and introduce students to video editing in a tangible way. The research had similar findings to Zigelbaum et al.: the tangible user interface increases collaboration and makes it easier to grasp abstract concepts, but the user interface lack efficiency and an easy way to undo actions.

Diogo Cabral and Nuno Correia explored pen-based video editing by creating a new graphical user interface and interaction style (Cabral and Correia 2017). Instead of just using the pen as a pointer device, the researchers designed an interaction technique they called "video as ink" where the user would "dip" the pen in a video clip and brush the content across a two-dimensional canvas to create sequences. This is an innovative take on video editing as the user interface bears little resemblance to either film editing or conventional video editing software. The research concludes that users considered the interaction to be attractive, creative and easy to use among other things, but only one of the twelve participants considered it efficient. This indicates that innovative and novel interaction techniques can be exciting and promote creativity, but these attributes are only a part of the user experience and especially for professional editors being fast and efficient is key. Familiarity should also not be overlooked as a positive attribute as this will mean less time spent on learning and more time spent on creating.

Several online posts exist on using video game controllers, drawing tablets and other input methods for video editing, but these are mostly anecdotal findings or overviews of the available products on the market. Edward Doherty tried a video game controller to prevent Repetitive Strain Injury and while the controller could not replace the keyboard and mouse for all tasks, it had enough functions available to be used for looking through the footage and the initial assembly (Doherty 2020). Video game controllers are certainly an interesting avenue as they minimize the need for hand movements and can keep the hands in a better position overall compared to the flat position on the keyboard and mouse. Jonny Elwyn compiled an overview of different options for editing control surfaces, starting from slightly customized keyboards and using additional buttons on a mouse for macros to software-specific control surfaces for certain tasks such as colour grading (Elwyn 2014). The level of customizability varies greatly, as does the intended use case and the problem each device is meant to solve. Some of the suggested control surfaces change the ergonomics of the interaction but the interaction technique and its goal remain the same. For example, a drawing tablet can replace the mouse as a pointing device but it simply takes over the tasks the mouse would be used for, not adding a new interaction technique or modality to the user experience.

2.5 Exploring the ISO Standard for Human-Computer Interaction

The ISO 9241 is a series of standards covering the Ergonomics of human-system interaction. Because this research is about the design of a physical control surface for human-computer interaction, these standards provide helpful definitions, principles and guidelines for the evaluation of existing control surfaces and the design process for a new concept. Applicable parts of the standard are reviewed in this chapter, focusing on the physical side of the interaction. ISO 9241-210:2019 describes the Human-Centred Design (HCD) process which is explored in detail in chapter 3.

ISO 9241-11:2018 contains the definitions and concepts for the standard series. Two of the most important definitions are usability and user experience. Usability is described as "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in the specified context of use". User experience is described as "user's perceptions and responses that result from the use and/or anticipated use of a system, product or service" (*ISO 9241-11:2018*).

ISO 9241-5:1998 describes the considerations for workstation layout and postural requirements. There are few applicable requirements in this part but there is one good recommendation: "Workplace design should be preceded by an analysis of the tasks that it is to support" (*ISO 9241-5:1998*).

ISO 9241-110:2020 covers the interaction principles which are meant for interaction design but are also applicable as heuristics for user interaction. The seven principles identified are suitability for the user's tasks, self-descriptiveness, conformity with user expectations, learnability, controllability, user error robustness and user engagement (*ISO 9241-110:2020*). Each principle is divided into subcategories and design recommendations are made for each category.

ISO 9241-112:2017 is focused on the presentation of information. Although this research is about primarily a user interface for input, the interface can still give feedback and present information, so this part of the standard is relevant. The principles for information presentation are detectability, freedom from distraction, discriminability, unambiguous interpretability, conciseness and consistency (both internal and external) (*ISO 9241-112:2017*). The standard is mostly concerned with visual information presentation but as general heuristics, the principles are applicable to other modalities as well.

ISO 9241-400:2007 presents the principles and requirements for physical input devices. The device-specific design requirements in the ISO 9241-4xx standards get a bit too technical for human-centred design and the standards predate the advent of some modern input techniques, such as capacitive touchscreens, but the generic design requirements are still valid. The generic requirements are appropriateness, operability, controllability and biomechanical load (*ISO 9241-400:2007*). These are further divided into sub-categories which are closely matched to the interaction principles in ISO-9241-110.

2.6 Summary

Modern non-linear video editing bears little similarity to its origins in film and linear video editing, even though the goal of the editor remains the same. The digital workflow also allows an editor to take on more responsibilities than ever before, but this research is focused on the editing phase, starting with ingestion and logging, going through the assembly and rough cut towards the iterative process of fine cutting and additional audio editing. The non-linear editors meant for complex editing are desktop software, most commonly used with keyboard and mouse even though historically editing is tied to specialized control surfaces. The user interfaces of modern NLEs are very similar and most often the use is centred around the timeline and the program output.

Previous research on video editing has been done around collaborative interaction and completely new paradigms for the graphical user interface, but most complex video editing is still the domain of expert users operating the software alone, with the director sometimes taking part in the creative decisions made during the process. There is little scientific research on the interaction methods used for video editing and most sources are anecdotal. The ISO-9241 series of standards offer valuable guidelines and principles for human-computer interaction in general, and they can be applied to video editing control surfaces as well.

3. OVERVIEW OF THE RESEARCH PROCESS

As the goal of the research is to come up with a set of design guidelines for new control surfaces for video editing, an effective way of achieving this is to go through a human-centred design process and evaluate both the process and the outcome. According to ISO 9241-210:2019, Human-Centred Design is "an approach to interactive systems development that aims to make systems usable and useful by focusing on the users, their needs and requirements" (*ISO 9241-210:2019*). The standard describes the principles that human-centred design should follow regardless of the process. The design should be based upon understanding of users, tasks and environments. Users include the end-users and other stakeholders affected by the use, either directly or indirectly. Users should be involved in the design and development process and the design should be driven by user-centred evaluation. The process should also be iterative and refinements should be equally user-driven.

There are some caveats to following the process defined by ISO 9241-210 in the context of this research. According to the standard, the design needs to address the whole user experience so peripheral parts of the system such as support, documentation, maintenance and product lifecycle should be considered. Due to the scope of the research the primary focus of the design process will be on the product itself and the user experience of the broader design will be considered when forming the guidelines. Another issue is that the design team should include multidisciplinary skills and perspectives. In the scope of this research inclusion of further design team members is not possible, so extra care should be taken when considering all of the perspectives of the design.

By setting UX and usability goals for the design, designing the product and evaluating the prototype, both the accuracy of the original design goals and the need for additional guidelines can be evaluated. The original design goals should be set based on user research, and existing products on the market can also inform decisions in the design process.

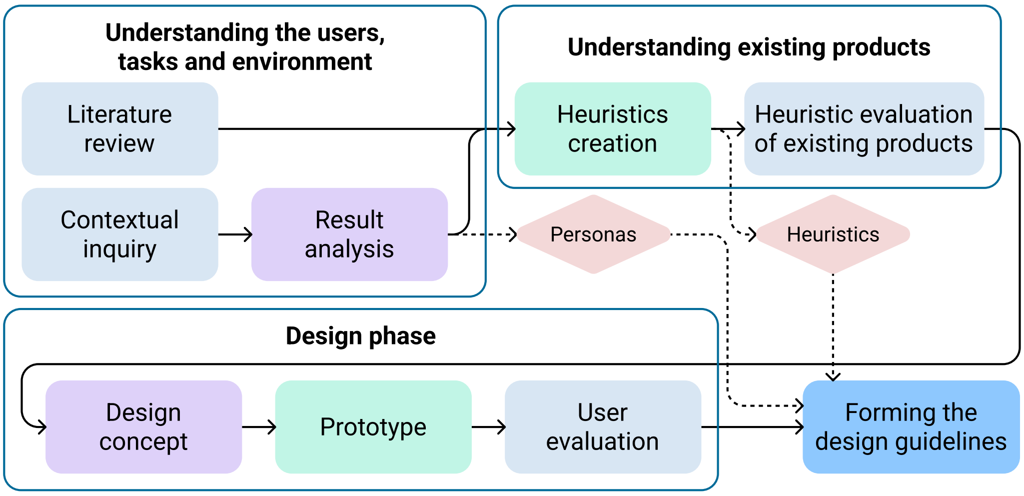


Figure 3.1. *Research process overview*

The individual phases of the process and methodology are discussed in more detail in the beginning of the chapters, but the entire process is depicted in figure 3.1 and summarized here. The process starts with a contextual inquiry to understand the user needs and the context of use. The results are analysed with an affinity diagram to form tentative user experience goals and personas for the product design, and potential heuristics to evaluate existing products. A set of heuristics for evaluating video editing control surfaces is created based on previous research and the results of the contextual inquiry, and a selection of existing control surfaces are evaluated using the heuristics to discover potential usability and user experience issues and ideas for a new concept. After the evaluation of existing products and solutions, a concept for the new product is formed. The concept is evaluated with a low fidelity prototyping technique. Finally, the design guidelines are formed based on results from the previous research phases. The methods for each phase are also evaluated after the phase is complete, and once more at the end of the research process.

4. CONTEXTUAL INQUIRY

Contextual inquiry is a common method for human-centred research for gathering information about the users, the environment of use and the tasks at the beginning of the research. Karen Holzbatt and Hugh Beyer discuss the principles of contextual inquiry in their book Contextual Design: Design for Life (Holtzblatt and Beyer 2016, pp. 43–73). The core of contextual inquiry is immersing the researchers in the end-users' life and context of use, and adopting a master-apprentice relationship where the researcher is the apprentice, learning about the tasks of the master as they occur. As the user is doing their tasks the researcher may ask questions and have discussions with the user. Sometimes these discussions lead to "retrospective accounts", where the user and researcher walk through past events related to the current tasks, thus expanding the timespan that the contextual inquiry covers.

The contextual inquiries consisted of an initial interview about the participants' backgrounds in video editing and their current work, an observation period where the participant worked on a video project, and a wrap-up interview with a broader set of questions about video editing and control surfaces. The structure is presented in appendix A.

Five participants were selected for the contextual inquiry. Originally the goal was for participants to be professional video editors but after three sessions with professionals the initial results suggested that amateur editors could also provide valuable insight and a different perspective to the research. For this reason the target group for this phase of the research was changed and the final two sessions were done with amateur editors. The professional statuses, backgrounds and used software of the participants are listed in table 4.1.

Table 4.1. *Contextual Inquiry Participants*

Pro	Experience	Experience with control surfaces	Editing software
Yes	Editing experience from a long period of time starting with tape editing and TV, focuses on documentary editing	Experience with linear and non-linear editing stations before computers, editing-specific keyboards	Final Cut Pro 7
Yes	Editing experience in commercial videos, moved to motion graphics	Experience with a gaming keypad and colour grading panels	Adobe Premiere Pro CC
Yes	Editing experience in commercial videos	Experience with colour grading panels	Adobe Premiere Pro CC
No	A few years of editing experience in short films and other short format videos	No experience	Blackmagic DaVinci Resolve
No	Editing experience as a freelance videographer in short format videos and theatre plays	Experience with Loupedeck CT	Adobe Premiere Pro CC

The process outlined by Holzbatt and Beyer also includes an interpretation session after the contextual inquiry sessions (Holtzblatt and Beyer 2016, pp. 81–84), but most of the process is meant to build a shared understanding for multi-member teams. The interpretation session should produce a set of notes collected from the contextual inquiry materials or insights from the researchers. This data is then analysed using an affinity diagram in which similar items are grouped together and form larger themes and commonalities (Holtzblatt and Beyer 2016, pp. 127–131) Notes were first grouped under coarse themes and the further grouped to find common opinions or ideas. Some of these groups were developed into tentative UX goals or heuristics for evaluating control surfaces. The affinity diagram is presented in appendix B and the results of the contextual inquiry are divided into the themes found during the analysis.

The results show how much variation there is between individuals regardless of whether they are professional or amateur editors. The work environments, hardware and software are a part of it, but more importantly personal workflow and the type of work the editors do vary greatly. The analysed results from the affinity diagram are divided by theme and some of the most important findings for each theme are described.

4.1 Environment

The work environments are dependent on the space they have to be built in and most often amateur or freelance editors have a space that needs to serve multiple purposes. The space is often a home office or a workstation that is also used for other tasks as well so significant modifications to the space or the workstation itself are often not feasible. Otherwise, the environments ranged from dimly lit personal office to bright an open multi-person office room, to a dedicated editing suite with acoustic treatment. Three of the five desks were electric ones and at two users stated working in a standing position occasionally, so that is something to keep in mind when considering the ergonomics of control surfaces. An example of a freelance editor's environment is presented in figure 4.1.

Figure 4.1. *Example of an editing workspace*

Professional workplaces can justify dedicated editing spaces and build them out to match the requirements, sometimes adding large purpose-built control surfaces for the editors. Despite this, of the five environments involved in the contextual inquiry none had editing-specific hardware permanently installed. One of the production companies that participated in this phase was quite large at 26 people, but they concentrated on small productions, mainly commercials and short corporate videos. This might partially explain the lack of specialized hardware even in their case, but a more likely explanation is simply not needing such equipment.

4.2 Hardware Tools

Editing-specific hardware has originated from the unique task it had to accomplish and the user interfaces were created based on the capabilities of the editing machines (Rubin and Diamond 2000, pp. 41–45). As the video industry moved towards computers, some specialized control surfaces were designed for editing, but keyboards started to become more and more common as a mode of input. The current generation of video editors have largely started their careers in video production during a time when there is no need to use an expensive control surface, instead relying on the keyboard and mouse as input devices. The combination is already familiar for new editors and with some practise keyboard shortcuts for the most common actions are committed to muscle memory, so these users might not feel the need for an editing-specific control surface at all.

Of the five participants in the study one had experience with pre-computer editing units, and he had also switched to using a keyboard and mouse for his current work. Only one participant was actively using an additional device but it was supplemental to keyboard and mouse, not able to replace either. A recurring theme with all participants was that video editing today includes a lot of other tasks than using the editing software, for example searching for stock material and music, so text input and pointer are likely unavoidable parts of the interaction in any case.

When it came to video editing all participants seemed to agree that muscle memory and an efficient way to use the most common actions are both important. Some participants had created their own keyboard shortcut maps for the programs they used while others had memorized the default shortcuts for their program. Based on the observation and subsequent interviews, the bottleneck of the process was in the creative problems, meaning the content and storytelling of the video, rather than the technical execution of each task.

4.3 Learning Curve and Adopting New Ways of Working

On the topic of introducing a new control surface to the workflow, one of the big concerns for both amateurs and professionals alike was the learning curve of the new device. Professional editors have to deliver finished products on time which puts pressure on them to get things done during work hours. Any time taken to learn a new control surface would be time away from productive work and according to one participant, the value proposal of the new control surface would have to be very clear and convincing to get users to take it into use. Some of the professional participants also mentioned not wanting to spend time outside of work on learning something new, so the learning would have to happen during work hours. They also had the hypothesis that amateur editors might have more interest in refining their workflow.

This theory seemed to be somewhat confirmed by the two amateur participants, one of whom was using a Loupedeck CT control surface and the other was seriously considering buying the first party control surface for their editing software. However, the latter also mentioned that time taken to learn the control surface would be less time to spend on making the videos themselves, but it might still be a lower threshold than for a professional. Both amateur and professional editors agreed that an editor who is just starting out might be the ideal candidate to take up a control surface as they aren't set in any habits yet. However, a beginner editor is also a very unlikely user for a specialized control surface due to the additional cost of obtaining one, and possibly lack of awareness about the existence of these devices.

4.4 Conclusion from the Contextual Inquiry

The contextual inquiry overturned some of the earlier hypotheses about willingness to learn and shifted the focus of the concept design as well. Originally the target audience of the research was professional video editors because of the hypothesis that professionals would be more willing adopt a new control surface than amateurs, thanks to the potential for more efficient workflow and time saving in the long term. During the contextual inquiry it became clear that this is not the case due to professional editors having stricter schedules than amateurs and productions they need to complete.

Because of this finding, amateur editors were included in the contextual inquiry and they provided valuable information about the differences between them and professionals. Both groups valued the time they could spend working on a video more than time spent learning a new way of working but amateur editors seemed more willing to adopt new devices and consider the time spent learning as a part of the hobby. Professional editors were not opposed to adopting a new control surface either, but for them the time investment seemed to be a bigger issue.

One of the most important takeaways from this phase is that the user needs and use cases vary between individuals and even between projects so much that the control surface has to be flexible to serve the needs of more than one person. Regardless of the user's experience level, it is also important that the adoption of the new control surface is effortless and the value proposal of the device is clear even before hands-on experience with it.

4.5 Products of the Contextual Inquiry

Because of the findings of the contextual inquiry, the target user group for the research was changed to both professional and amateur editors, who are interested in improving their workflow and willing to adopt new ways of working.

Tentative UX goals were defined for the product design:

- Sense of freedom and control: feeling like you can make the product do what you need it to do and fit it to you personal workflow
- Satisfaction: both the product itself and its use give the user a feeling of satisfaction and accomplishment
- Unobtrusiveness: both starting the use and the continued use do not cause the user issues, instead reducing the barrier between creative thinking and execution

Two different personas were also created to represent the archetypes of editors and aid in the design process. Personas are highly specific descriptions of a typical user of the product that help the designers emphatize with the end users better than with an abstract description (Harley 2015). The first persona, presented in figure 4.2, represents a serious hobbyist who likes to develop the process of editing in addition to enjoying the end result. The second persona, presented in figure 4.3, represents a professional editor who prefers to use their time on editing and focus on delivering the end product.

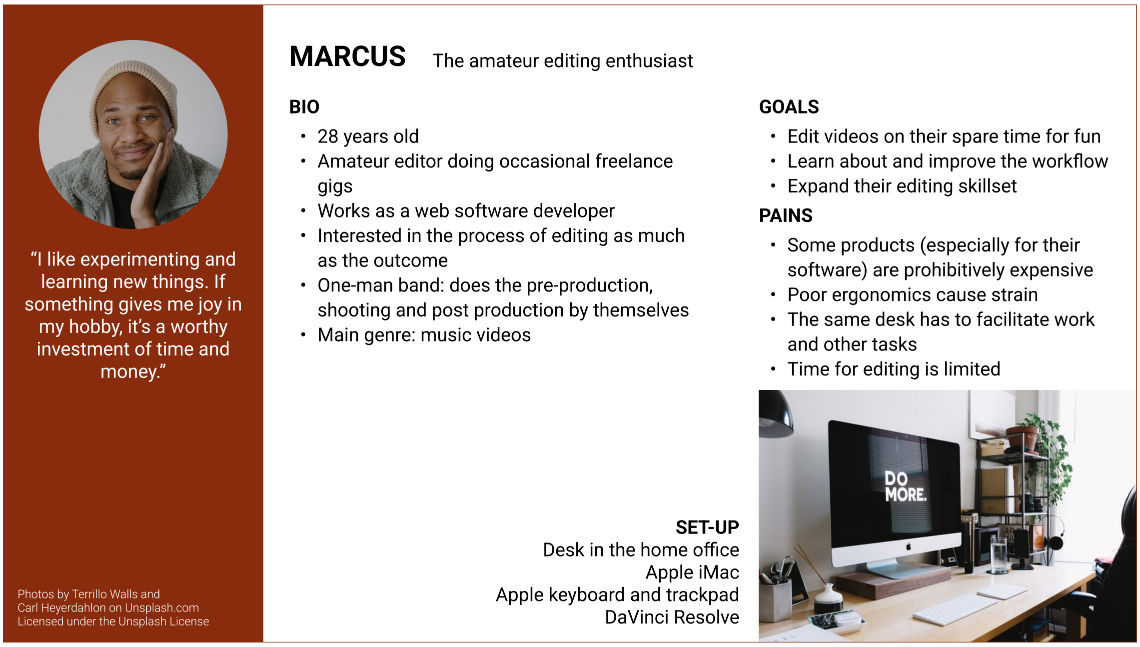


Figure 4.2. *Persona 1: Marcus*

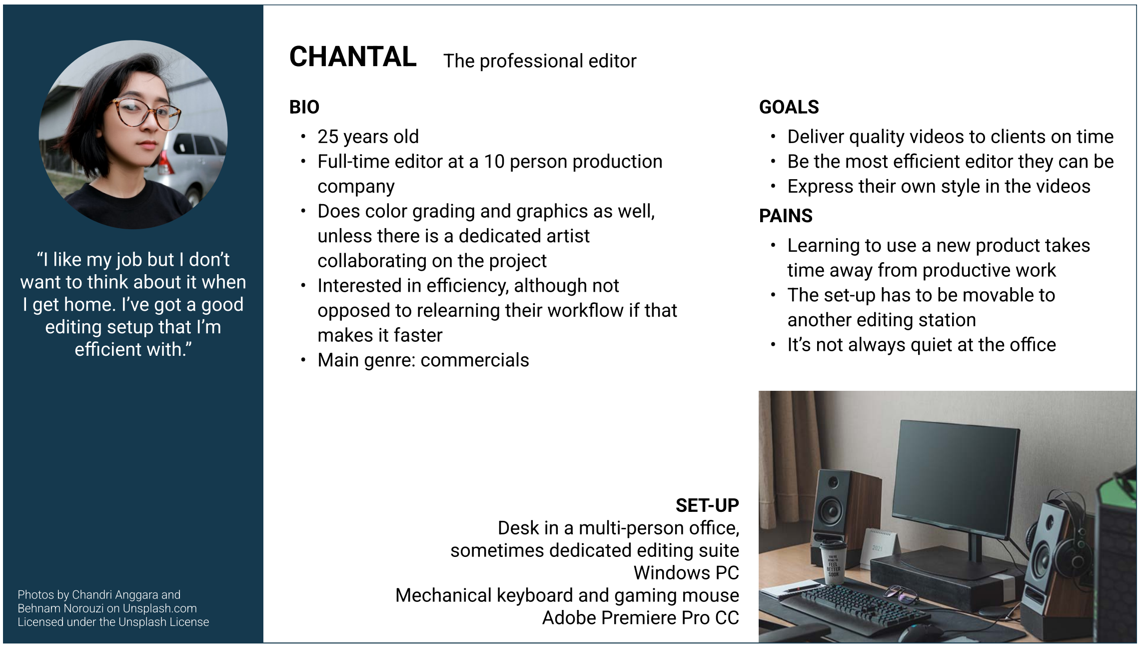


Figure 4.3. *Persona 2: Chantal*

5. HEURISTIC EVALUATION OF EXISTING CONTROL SURFACES

Heuristic evaluation is a technique for usability and user experience evaluation based on heuristics, a set of principles or general rules created for this purpose. The goal of the method is to identify usability problems, but it can also highlight the good parts of a product. Heuristic evaluations are conducted by usability experts, by exploring and using the user interface and evaluating it against the heuristics. (Nielsen 1994)

A selection of commercially available control surfaces were selected to be evaluated, but few heuristic sets exist for evaluating physical user interfaces and none match the application domain of video editing well. Because of this a method for creating usability and user experience heuristics designed by Quiñones et al. (Quiñones, C. Rusu and V. Rusu 2018) was used to develop and refine new heuristics.

5.1 Forming the User Experience Heuristics

The formalized method for creating heuristics consists of 8 stages beginning with literature review and other information gathering, continues into forming the heuristics and ends with refining and evaluating the created heuristics.

The method is meant to be iterative and the definition, refinement and evaluation of heuristics is meant to be repeated until a set of heuristics is found that adequately evaluates the application domain. In this research, the validation and refinement stages were done as a part of the specification stage due to the limited scope. Expert judgement was used for the validation in this stage, and a more comprehensive evaluation of the method for forming the heuristics as well as the created heuristics is presented in section 5.3.

The original research by Quiñones et al. seems to focus on evaluating software, so some of the terminology of the described process has been changed to reflect the nature of the evaluated system. The follow-up research on applying the method to create user experience heuristics (Quiñones and C. Rusu 2019) was utilized to guide each stage of the process.

5.1.1 Exploratory and Experimental stage

The first stage of the method is exploratory stage, focusing on gathering information about the application domain via literature review. The second stage is the experimental stage and it expands on the first stage's results. Possible experiments include for example heuristic evaluation with existing heuristics, surveys and interviews. In this research the first stage was completed during the initial literature review, revealing that there is little research done in the domain of video editing, and most of the information is anecdotal reports by individuals working with video, but not necessarily researchers themselves. General definitions and principles related to physical user interfaces were collected. The contextual inquiry was done as a part of the experimental stage and in addition multiple editing sessions were done on a video editing project to gather further information.

The first product of the exploratory stage is information about the system, which includes relevant definitions of the domain, and general and specific features of the system. The general features in this case are related to physical user interfaces and the specific features are features of editing control surfaces. The exploratory stage should also produce relevant usability and UX attributes, most often selected from literature, and existing sets of heuristics or other relevant items. Some of the common tasks in video editing were collected in this stage.

The definitions of the system, presented in table 5.1 are mainly generic definitions for user experience, usability and other relevant topics. The definitions for video editing and video editing control surfaces were created during this research while other definitions come from the ISO 9241 series of standards.

Table 5.1. *Definitions*

Information type	**Description**
Definition of video editing	Defined in section 2.1: assembling media to a sequence to form a cohesive piece
Definition of video editing control surface	Defined in chapter 1: a physical user interface used for the task of video editing, containing the hardware and the software that enables the interface to function with existing editing software.
Definition of usability (ISO 9241-11)	Extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in the specified context of use (*ISO 9241-11:2018*)
Definition of user experience (ISO 9241-11)	User's perceptions and responses that result from the user and/or anticipated use of a system, product or service (*ISO 9241-11:2018*)

The features of the system, split into general and specific features, are presented in Tables table 5.2 and table 5.3. These features are collated from standards, related research and the contextual inquiry.

Table 5.2. *Generic features of physical user interfaces*

Feature	Description
Ergonomics	Understanding of interactions among human and other elements of a system, optimizing human well-being and overall system performance (*ISO 9241-11:2018*). This includes efficiency of motion, considering biomechanical load, facilitating different working positions (also part of accessibility) and avoidance of physical harm
Modalities	The channels of sensory input and output used by the user interface. (*ISO 9241-112:2017*)
Interaction types	The ways in which a user interacts with the system, defined as five types: instructing, conversing, manipulating, exploring and responding (Sharp 2019).
Feedback	The user interface can present information about the software state and the state of the interface itself
Accessibility	Extent to which systems can be used by people from a population with the widest range of user needs (*ISO 9241-11:2018*)
Recognition and recall	Recognition: identifying specific parts of the user interface (and the tasks facilitated by them) based on cues provided by the user interface and the context. Recall: identifying specific parts of the user interface by remembering the previous use of the system, with few or no cues provided by the user interface (Budiu 2014). Muscle memory relies on recognition as well since the user needs to recognize the current state of the user interface and locate the action they need, but the recognition often moves from visual to tactile cues.
Form factor	The possible environments of use depend on the size and form factor of the physical interface. Some user interfaces are non-movable and need to be permanently installed, while others can be stored away when needed.

Continued on the next page

Table 5.2 – *Continued from previous page*

Feature	**Description**
Learnability and ease of setup	Learnability considers the ease of starting to use a new system, the time it takes to become efficient with it and the ultimate plateau of efficiency of use once the users have fully learned to use it (Joyce 2019). In addition, the initial effort to set up the system before the use, as well as any setup that needs to be done between uses should be considered.
Degree of specificity	Level of specialization of the user interface towards a single task. A high degree of specificity may increase efficiency in the selected task but lower the user interface's versatility and efficiency in other tasks.
Consistency of action	A single action on the user interface causing the same effect regardless of the context. This helps reduce the cognitive load of the user interface but may decrease the amount of tasks a user can accomplish with the interface.
Individualization	Modification of interaction and presentation of information to suit individual capabilities and needs of users (*ISO 9241-171:2008*)
Hedonic attributes	Aesthetics: professional look, feel of quality, fun factor. Attributes which do not contribute to the fulfilment of the user's goals, but promote psychological wellbeing and contribute to the user experience (Hassenzahl 2004)

Table 5.3. *Specific features of video editing control surfaces*

Feature	Description
Integration with editing software	Level of access to the features of the software, for example actions easily accessible through keyboard shortcuts requires low level integration, while actions not available through other means than navigating multiple menus made available on the control surface could be considered high level integration
Relationship between interaction and the user interface	The resemblance of the user interface to what controls are displayed and used on the software UI can wary depending on the control surface, but generally control surfaces that closely resemble the UI of the task they are designed for are considered useful
Companion software	The control surface may require a companion software to be installed on the computer it is used with, often used for individualization and providing a higher level of integration with the NLE and other software

The domain-specific tasks of video editing are presented in table 5.4. This list is not exhaustive, but is a good overview of the different stages of the workflow and different actions available to the editor.

Table 5.4. *Domain-specific tasks*

Task	Description
Navigating panels	Focusing input on different panels in the user interface
Browsing raw material	Project bin navigation, selecting material for preview
Navigating the timeline	Zooming and panning in horizontal and vertical directions (time and tracks)
Jogging	Navigating the time axis of the timeline or a single clip while playing the material, moving the current time indicator
Inserting to timeline	Setting in and out points on a clip and inserting it to the position of the current time indicator or an arbitrary position
Clip selection	Selecting single or multiple clips, choosing the tracks from which clips can be selected
Clip manipulation	Move, cut, delete, trim head or tail, slip (changing the in and out points of the source material without changing the position or duration of the clip)
Cut manipulation	Move the cut point of two clips without changing the position of the clips
Clip attribute manipulation	Adding or modifying keyframes to control a property of the clip, such as video opacity or audio level

The usability/UX factors consist of usability attributes, principles from the ISO 9241 series standards and existing heuristic sets collected during the literature review:

1. Nielsen's usability attributes (Nielsen 2012)
2. Peter Morville's UX attributes (Morville 2004)
3. Interaction principles from ISO 9241-110:2020 (*ISO 9241-110:2020*)
4. Information presentation principles from ISO 9241-112:2017 (*ISO 9241-112:2017*)
5. Generic requirements for physical user interfaces from ISO 9241-400 (*ISO 9241-400:2007*)
6. Nielsen's ten heuristics (Nielsen 2020)
7. Heuristics for physical environment by Shanklin et al. (Shanklin et al. 2020)

Two sets of heuristics were selected at this stage: Nielsen's heuristics and the heuristic set created by Shanklin et al. to evaluate physical environments, specifically rail transit systems (Shanklin et al. 2020). The experimental stage should yield problems with existing heuristics and there were some issues found with Nielsen's heuristics, as they are more focused on evaluating graphical user interfaces and not all of the heuristics are applicable to video editing:

2. **Match between the system and real world.** There is no good real world equivalent to the video strip except for film, and very few people have experience with physical film cutting, so the match is not useful.
6. **Recognition rather than recall.** The meaning of this heuristic changes as the user's proficiency with the task increases. At first recognition is an important part of helping the user find the actions on the control surface but as the user gains experience and builds muscle memory, recognition becomes more about recognizing the current state of the user interface so that the user can make the correct actions to complete their task.

The heuristics created by Shanklin et al. are based on Nielsen's heuristics with 5 of the 11 heuristics being newly created by the research group. These heuristics are more focused on a user's journey through a physical service like public transport, but are mostly applicable to physical user interfaces as well.

The last product of the experimental stage is potential usability or user experience problems at this stage. While there are no definite problems, some potential issues do exist with using keyboard and mouse as the input devices: the efficiency of the input methods might not be ideal, building the muscle memory for the most common actions takes time, and especially the keyboard offers very little feedback about the hand's position on it, so erroneous actions are possible.

5.1.2 Descriptive stage

In this stage the products of the previous stages are prioritized within each topic on a number scale: 1 is not important, 2 is somewhat important and 3 is highly important. The method does not define any criteria for the prioritization, but the goal is to select information relevant to the creation of new heuristics. All items were considered highly important by default and justification for lowering the priority is given. Items with a priority of 1 are discarded from the next stages. The products of the descriptive stage is selected information about the application, selected features of the application domain, selected usability/UX attributes, selected sets of heuristics and other relevant items.

The prioritized definitions are presented in table 5.5. All definitions from the exploratory stage were considered highly important.

Table 5.5. *Prioritized definitions*

Value	Information type	Description
3	Definition of video editing	Defined in section 2.1: assembling media to a sequence to form a cohesive piece
3	Definition of video editing control surface	Defined in chapter 1: a physical user interface used for the task of video editing, containing the hardware and the software that enables the interface to function with existing editing software.
3	Definition of usability (ISO 9241-11)	Extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in the specified context of use (*ISO 9241-11:2018*)
3	Definition of user experience (ISO 9241-11)	User's perceptions and responses that result from the user and/or anticipated use of a system, product or service (*ISO 9241-11:2018*)

The prioritized features of the domain are presented in table 5.6. While no features were outright discarded, some were considered less important than others.

Table 5.6. *Prioritized features of the domain*

Value	Information type	Justification for lower priority
3	Ergonomics	
3	Modalities	
3	Interaction types	

Continued on the next page

Table 5.6 – *Continued from previous page*

Value	Information type	Justification for lower priority
2	Feedback	The NLE provides visual and auditory feedback about the state of the software and the sequence. The control surface should provide feedback on the state of the control surface itself and it may augment the feedback from the NLE, but the latter is not required
2	Accessibility	Video editing is limited in terms of accessibility, the control surface should take a wide range of users into consideration but limit the allowances to users capable of using the editing software and capable of the task of video editing
3	Recognition and recall	
2	Form factor	Large size or requirement of permanent installation may rule out some users but that can be acceptable if the target user group is very specific
3	Learnability and ease of setup	
3	Degree of specificity	
3	Consistency of action	In different contexts this can be prioritized differently, for example an otherwise unused control can be utilized for another action when switching modes from assembly to editing but globally applicable actions should behave in the same way regardless of the context
2	Individualization	A well-designed control surface with no individualization options can become efficient if the user is willing to adapt to the new way of working
2	Hedonic attributes	The control surface is ultimately a tool and pragmatic attributes outweigh the hedonic

Continued on the next page

Table 5.6 – *Continued from previous page*

Value	Information type	Justification for lower priority
2	Integration with editing software	A high level of integration is not necessarily required for good usability and user experience
2	Relationship between interaction and the user interface	Replicating the software user interface may only serve very specific tasks or the control surface becomes big and complex. The degree of specificity also increases
2	Companion software	Ultimately the goal is to set the device up once and the repeated use concentrates on interacting with the device itself

The usability and UX attributes gathered in the exploratory stage had many overlaps so some of the attributes were discarded for that reason. The prioritized attributes are presented in table 5.7.

Table 5.7. *Prioritized usability/UX attributes*

Value	Attribute	Author	Justification for lower priority
3	Learnability	Nielsen	
3	Efficiency		
3	Memorability		
3	Errors		
3	Satisfaction		
3	Useful	Morville	
3	Usable		
2	Desirable		
1	Findable		Not well applicable to the physical context
2	Accessible		
3	Credible		
3	Valuable		

Continued on the next page

Table 5.7 – *Continued from previous page*

Value	Attribute	Author	Justification for lower priority
1	Effectiveness	ISO 9241-11	The Nielsen attributes are similar and more complete, so the ISO set is discarded
1	Efficiency		
1	Satisfaction		
ISO 9241 principles			
3	Suitability for the user's tasks	ISO 9241-110	
2	Self-descriptiveness		A certain level of initial learning can be accepted for a product meant for expert use
3	Conformity with user expectations		
3	Learnability		
3	Controllability		
3	User error robustness		
3	User engagement		
2	Detectability	ISO 9241-112	The information the control surface will present is secondary to the information of the NLE user interface
3	Freedom from distraction		
3	Discriminability		
3	Unambiguous interpretability		
3	Conciseness and consistency (internal and external)		
3	Appropriateness	ISO 9241-400	
3	Operability		
1	Controllability		Duplicate

Continued on the next page

Table 5.7 – *Continued from previous page*

Value	Attribute	Author	Justification for lower priority
3	Biomechanical load		

For the heuristics, all of Nielsen's heuristics were included for but as the heuristics by Shanklin et al. were based on them, only the new heuristics were included for prioritization. The prioritized heuristics are presented in table 5.8.

Table 5.8. *Prioritized heuristics*

Value	Attribute	Author	Justification for lower priority
3	Visibility of system status	Nielsen	
3	Match between system and the real world		
3	User control and freedom		
3	Consistency and standards		
3	Error prevention		
2	Recognition rather than recall		The importance and mode of recognition changes over repeated use
3	Flexibility and efficiency of use		
2	Aesthetic and minimalist design		The focus of the control surface is primarily in its function, although minimalistic design and avoiding overcomplicating the user interface is still important
3	Error recovery		
3	Help and documentation		
3	Human limitations	Shanklin et al.	
3	Predictability		

Continued on the next page

Table 5.8 – *Continued from previous page*

Value	Attribute	Author	Justification for lower priority
3	Accommodation		
3	Accuracy		
3	Suitable tempo		

5.1.3 Correlational stage

The purpose of the correlational stage is to match features, usability/UX attributes, heuristics and other relevant elements together. If no matching heuristics are found for a feature, a new heuristic can be proposed. The products of this stage are matched features, attributes, existing heuristics and other relevant items. The correlated items are presented in table 5.9 and in many cases, the matches overlap in some of the categories. The final heuristics will not be isolated entities and instead many heuristics will be interconnected.

Table 5.9. *Matches of control surface features to usability/UX attributes or other relevant items and existing heuristics*

Feature	Attribute or relevant item	Heuristic
Ergonomics	Efficiency, Operability, Biomechanical load	Human limitations
Modalities	Suitability for the user's tasks, Accessible, Usable, Desirable	Match between system and real world
Interaction types	Efficiency, Accessible, Controllability	Match between system and real world
Feedback	Errors, Self-descriptiveness, Detectability, Freedom from distraction, Discriminability, Unambiguous interpretability, Conciseness and consistency	Visibility of system status, Accuracy
Accessibility	Usable, Accessible, Controllability	Accommodation
Recognition and recall	Learnability, Memorability, Efficiency	Recognition rather than recall

Continued on the next page

Table 5.9 – *Continued from previous page*

Feature	**Attribute**	**Heuristic**
Physical attributes	Satisfaction, Desirable	Aesthetic and minimalist design
Learnability and ease of setup	Learnability, Errors, User error robustness	Error prevention, Error recovery, Help and documentation
Degree of specificity	Efficiency, Learnability	Flexibility and efficiency of use
Consistency of action	Memorability, Errors	Consistency and standards, Predictability
Individualization	Controllability, User engagement	Flexibility and efficiency of use, User control and freedom, Accommodation
Hedonic attributes	Satisfaction, Desirable	Aesthetic and minimalist design
Integration with editing software	Efficiency, Useful, Valuable, Controllability	User control and freedom
Relationship between interaction and the user interface	Conformity with user expectations, Appropriateness	Match between system and real world
Companion software	Learnability, Efficiency, Usable	Flexibility and efficiency of use, Accommodation

5.1.4 Selection stage

The purpose of the selection stage is to identify which existing heuristics to keep, adapt or eliminate, and if any new heuristics need to be created. The heuristics are also ranked on the scale of useful (1), important (2) and critical (3). Heuristics with N-prefixed IDs are from Nielsen, S-prefixed heuristics from Shanklin and heuristics without an ID are new heuristics created for evaluating control surfaces. The adapted heuristics were handled in three different ways: A single new heuristic matches the adapted heuristic, multiple new heuristics cover parts of the adapted heuristic, or multiple adapted heuristics were fused into a single new heuristic. The heuristics and the planned actions are presented in table 5.10.

Table 5.10. *Heuristic selection process*

ID	Heuristic name	Action	Feature covered	Priority
N1	Visibility of system status	Adapt	Feedback	1
N2	Match between the system and the real world	Adapt	Interaction types, Relationship between interaction and the user interface	1
N3	User control and freedom	Adapt	Individualization, Integration with editing software	3
N4	Consistency and standards	Adapt	Consistency of action	2
N5	Error prevention	Adapt	Learnability and ease of setup	3
N6	Recognition rather than recall	Adapt	Recognition and recall	3
N7	Flexibility and efficiency of use	Adapt	Degree of specificity, Individualization, Companion software	3
N8	Aesthetic and minimalist design	Adapt	Hedonic attributes	1
N9	Error recovery	Eliminate	-	-
N10	Help and documentation	Eliminate	-	-
S1	Human limitations	Adapt	Ergonomics	2
S5	Predictability	Adapt	Consistency of action	3
S6	Accommodation	Adapt	Accessibility, Individualization, Companion software	1
S8	Accuracy	Adapt	Feedback	3
S10	Suitable tempo	Eliminate	-	-
-	Learnability	Create	Learnability and ease of setup	3
-	Co-operation with other input devices	Create	Physical attributes	3

5.1.5 Specification stage

The specification stage formalizes the definitions of all heuristics with a priority, description, examples and a checklist for evaluation. The proposed heuristics are presented in table 5.11 and described in more detail in their individual tables. The heuristics are also ranked on the scale of useful (1), important (2) and critical (3).

Table 5.11. *Proposed heuristics*

ID	Name	Priority	Basis
CS1	Presentation of feedback and system status	2	N1, S8
CS2	Representative interaction	2	N2
CS3	Learnability	2	N5, N9, N10
CS4	Memorability and recognition	3	N6
CS5	Predictability and consistency	3	N4, S5
CS6	Flexibility and efficiency of use	3	N7
CS7	Individualization and user freedom	2	N3, S6
CS8	Error prevention	2	N5
CS9	Sustainable ergonomics	3	S1
CS10	Fit and finish	1	N8
CS11	Co-operation with other input devices	3	S6

The method created by Quiñones et al. suggests a template for defining the heuristics which is used in this stage. They also recommend developing a checklist for evaluating each heuristic as the heuristic is a general rule that can be difficult to evaluate. (Quiñones, C. Rusu and V. Rusu 2018) Control surfaces for video editing are varied in their form and function, so the checklists from the original template are modified to evaluate a wide variety of products: some of the items on the checklists are alternative, meaning one of the items from a set has to be checked, but there are different options to cover different control surfaces. It is not expected of any control surface to check every item on every heuristic, but checking at least one item on each heuristic is indicative of a good design.

Table 5.12. *CS1: Presentation of feedback and system status*

ID	**CS1**
Name	Presentation of feedback and system status
Priority	(2) Important
Definition	The control surface communicates information about the state of the NLE and the state of the control surface itself in a clear and understandable way.
Explanation	The control surface may communicate in a limited fashion about the state of the NLE that augments the graphical user interface of the computer. In addition, the control surface may have different operating modes or states of its own and the device must clearly communicate its current state to the user.
Examples	Macro keypad indicates which actions are currently available on the buttons, A motorized slider corresponding to a control on the user interface moves to match the position of the software.
Benefits	By providing multimodal feedback about the system status the user can more efficiently determine the state, what actions are available and how they can advance their task. Providing information about the state of the control surface lowers the risk of errors due to the user not remembering the state.
Checklist	1. If the control surface has multiple states, it communicates the state to the user 2. The control surface communicates information about the NLE state to the user
Feature	Feedback (generic feature)
Usability attributes and principles	Errors, Detectability, Unambiguous interpretability
UX attributes	Usable
Related Heuristics	N1, S8

Table 5.13. *CS2: Representative interaction*

ID	**CS2**
Name	Representative interaction
Priority	(2) Important
Definition	The control surface uses interaction elements that relate to the resulting action in the NLE in some way.
Explanation	The physical controls of the control surface are similar to the affected controls of the NLE, the interaction is representative of a physical manipulation of the user interface elements, or the interaction is relatable to the resulting action in other ways.
Examples	A real slider used to control a virtual slider (physical control - affected control), dragging interaction represents the real world (interaction - real world), navigation is done by rotating with a rotational control in the direction of the intended movement (relatable interaction).
Benefits	A clear connection between the physical action and the effect on the software allows more accurate and natural control for the user, and lowers the chance of moving a control in the wrong direction.
Checklist	1. Physical controls match the virtual controls of the NLE user interface 2. Controls are selected so that they are familiar to the user in the context of the action (e.g. slider for volume) 3. Interactions relate to the physical world where possible
Feature	Modalities (generic feature), interaction type (generic feature)
Usability attributes and principles	Learnability, Suitability for the user's tasks, Appropriateness
UX attributes	Accessible
Related Heuristics	N2

Table 5.14. *CS3: Learnability*

ID	**CS3**
Name	Learnability
Priority	(2) Important
Definition	The initial setup and first use of the control surface is fast and easy, and developing expertise with the control surface is made as easy as possible.
Explanation	For both inexperienced and professional editors, the initial learning process from unboxing to effective use needs to be quick. The development from novice user to expert who can make use of the advanced features of the control surface must be possible with little effort in addition to using the control surface for its intended purpose.
Examples	Good quick start guide, ability to adopt some of the features of the control surface at first and expand over time.
Benefits	Adopting a new control surface to the editing workflow should have a clear value proposition: the learning will not take away time from productive work and the continued use will increase productivity in the long term.
Checklist	1. The control surface comes with clear documentation for getting started 2. From the moment of unboxing, the first meaningful use of the control surface happens within an hour 3. A clear path for developing skills and improving expertise with the control surface is provided
Feature	Learnability and ease of setup (generic feature)
Usability attributes and principles	Learnability, User error robustness
UX attributes	Usable
Related Heuristics	N5, N9, N10

Table 5.15. *CS4: Memorability and recognition*

ID	**CS4**
Name	Memorability and recognition
Priority	(3) Critical
Definition	The control surface has features which allow the user to easily recognize its state, build muscle memory for repeated actions and re-establish proficiency after a pause in the use.
Explanation	The user's attention is primarily on the visual output of the NLE and effort should be directed to minimizing the need for the user to direct their attention to the control surface. This is mainly true for gaze and hearing, so other modalities can be used to communicate about the control surface in an unobtrusive way. The control surface should also be memorizable in a way that it's easy to start the use again after a period of not using it.
Examples	Keyboard home position indicators, colour coded keys, other tactile features which increase recognition.
Benefits	Being able to keep the user's attention on the video they are creating instead of the minutiae of how they are accomplishing each task increases the chance of maintaining a "flow" state while editing, ultimately increasing productivity.
Checklist	1. The user interface provides unobtrusive hints of the users position on it, allowing the user to recognize the actions they are searching 2. The control surface allows users to memorize certain actions or features and execute them consistently
Feature	Recognition and recall (generic feature)
Usability attributes and principles	Efficiency, Memorability
UX attributes	Useful
Related Heuristics	N6

Table 5.16. *CS5: Predictability and consistency*

ID	**CS5**
Name	Predictability and consistency
Priority	(3) Critical
Definition	A single action on the control surface always results in the same outcome or a similar one in the NLE, and the outcome is deducible from the state of the NLE and the control surface.
Explanation	A single action on the control surface results in similar outcomes in different contexts on the program UI where applicable, such as having different panels active. By looking at the state of the NLE and the feedback the controller gives, the user can accurately predict what each action will result in and choose the action that leads to their intended goal.
Examples	Jogging control always acts as a jogging control when a clip or a timeline is selected, if a clip is not selected the control may change zoom or pan. When jogging, the control will always move the same amount.
Benefits	Consistent behaviour reduces the chance of errors and user's frustration.
Checklist	1. Each action on the control surface always has the same result 2. If a control's role is dependent on the NLE or the state of the control surface, the state and outcome is clearly communicated
Feature	Consistency of action (general feature)
Usability attributes and principles	Memorability, Errors
UX attributes	Credible
Related Heuristics	N4, S5

Table 5.17. *CS6: Flexibility and efficiency of use*

ID	**CS6**
Name	Flexibility and efficiency of use
Priority	(3) Critical
Definition	The control surface caters to both inexperienced and expert users, allowing each user to work at their pace and level of complexity.
Explanation	Shortcuts meant for expert users can accelerate the work and provide a higher plateau of efficiency once the user learns the system comprehensively. The control surface should provide novice users with means to complete their tasks but also allow the users to continue learning and improving their workflow.
Examples	Mouse has simple actions available (use cursor as razor to split clip) but keyboard shortcuts can accelerate the tasks (Control + K = split clip at playhead).
Benefits	Allowing the users to work at their pace broadens the target user group and allows a user to start the use without too much learning, and develop advanced skills over time.
Checklist	1. The control surface has basic functionalities easily available 2. The control surface offers shortcuts for advanced users, but their use is not required
Feature	Degree of specificity (general feature)
Usability attributes and principles	Efficiency, Learnability
UX attributes	Valuable
Related Heuristics	N7

Table 5.18. *CS7: Individualization and user freedom*

ID	**CS7**
Name	Individualization and user freedom
Priority	(2) Important
Definition	The control surface allows customization of the interaction and actions to suit each users' workflow.
Explanation	Editing workflow varies between users and setups and the control surface should accommodate different ways of working.
Examples	Adjustable keyboard shortcuts, a companion software for customization.
Benefits	By allowing the user to continue to work in a familiar way and augment the workflow with new controls, the threshold for adopting a new control surface is lowered.
Checklist	1. The actions of the control surface can be changed to user preference 2. Available actions that can be assigned to the control surface cover a broad range of approaches and workflows
Feature	Individualization (generic feature)
Usability attributes and principles	Controllability, User engagement
UX attributes	Accessible
Related Heuristics	N3, S6

Table 5.19. *CS8: Error prevention*

ID	**CS8**
Name	Error prevention
Priority	(2) Important
Definition	The feedback from the control surface prevent the user from making errors.
Explanation	Some actions of the control surface may be context-dependent and the control surface should provide enough feedback to the user to prevent triggering unwanted actions. Error recovery is usually done via the undo-action of the NLE.
Examples	Displays on the control surface describing a control, control only being used for a single action.
Benefits	Error prevention and recovery are some of the most important features of exploratory interaction as they allow the user to try different things without too much consequence, and steer the user towards the right actions where possible.
Checklist	1. The control surface has indicators for what action a control will execute in a given context 2. The control surface only uses controls for specific actions that do not change 3. The overall design prevents errors, such as unwanted tap or press gestures 4. The control surface provides a way for error recovery
Feature	Learnability and ease of setup
Usability attributes and principles	Learnability, Errors, User error robustness
UX attributes	Usable
Related Heuristics	N5

Table 5.20. *CS9: Sustainable ergonomics*

ID	**CS9**
Name	Sustainable ergonomics
Priority	(3) Critical
Definition	The physical attributes of the control surface accommodate a variety of users and working positions, and using the control surface for extended periods of time does not produce strain on the body.
Explanation	The control surface should attempt to be accessible to a wide range of users, regardless of body type or some physical disabilities, and also support sitting, standing and other body positions. The sustained usage should not subject the user to physical strain or increase the risk of repetitive strain injury (RSI).
Examples	Natural hand position while using the control surface, ability to rest arm on the desk or an armrest during use.
Benefits	The only way to make repeated long-term use possible is sustainable ergonomics.
Checklist	1. The user can adopt different body positions while using the control surface 2. The body position required for using the device is resting and does not require constant muscle activation 3. The control surface is accessible to people with tactile or motor disability
Feature	Ergonomics
Usability attributes and principles	Efficiency, Operability, Biomechanical load
UX attributes	Usable, Accessible
Related Heuristics	S1

Table 5.21. *CS10: Fit and finish*

ID	**CS10**
Name	Fit and finish
Priority	(1) Useful
Definition	The control surface looks and feels appropriate for its intended purpose, is durable and maintainable.
Explanation	The control surface is primarily a tool for expert or professional work and hedonic qualities are a part of the impression it gives. Robustness and quality finish are important, while the overall design and aesthetic qualities are subjective but also contribute to the look and feel of the product.
Examples	Smooth rotary controls, dense structure, mouse can withstand the weight of a resting hand.
Benefits	The fit and finish contribute to the overall user experience and sense of satisfaction that comes from the use of the product itself.
Checklist	1. The control surface feels robust and has a durable construction 2. The product looks professional 3. Maintainability has been taken into account in the design
Feature	Physical attributes
Usability attributes and principles	Satisfaction, Conformity with user expectations
UX attributes	Desirable
Related Heuristics	N8

Table 5.22. *CS11: Co-operation with other input devices*

ID	**CS11**
Name	Co-operation with other input devices
Priority	(3) Critical
Definition	The adoption of a new control surface must not be more difficult for any task than using existing user interfaces.
Explanation	The control surface accommodates the use of text and pointer input devices for other tasks or replaces them and provides the same functionality. Video editors often change context during the workday to different software which requires basic text and cursor input (such as searching for stock assets) or need those input methods in the NLE, so any control surface should either provide those functionalities to the user or have such shape and size that a separate keyboard and mouse can be used without additional effort.
Examples	Small control surface next to keyboard and mouse, programmable keyboard with macro functionality.
Benefits	The user is more likely to adopt a new control surface if there are no blatant drawbacks.
Checklist	1. The control surface is small enough to be used alongside a text input and pointer device 2. The control surface provides text input or pointer interaction
Feature	Physical attributes
Usability attributes and principles	Suitability for the user's tasks
UX attributes	Usable, Valuable
Related Heuristics	S6

5.2 Evaluation of Existing Products

Three existing products were selected for the heuristic evaluation: Loupedeck CT, Tunks Ergo33 and Wacom Intuos Pro M. The heuristic evaluations were done in the form of an editing session, starting from the setup and first use of the products until the user interface had been explored sufficiently. The products were evaluated by using them to edit an existing video project that had most of the assembly cut done, so the use was focused on rough editing, fine editing, some transitional effects and audio editing.

Each product is introduced, the general interaction style and available actions are explained and some impressions on the advantages and disadvantages of the control surface are described. The heuristic checklist is also presented for each product, but it is important to remember that the checklist has very limited granularity, and two products might fulfil a particular criterion in different ways and to different degrees.

5.2.1 Loupedeck CT

Loupedeck CT is an editing control surface with multiple control types and integrations with different programs. The control surface is designed to be used with a keyboard and mouse, supplementing the keyboard with dedicated shortcuts, rotary controls and touch screen buttons for different tasks. Loupedeck CT offers a good mix of customizable buttons and rotary controls, as well as visual indicators for the controls whose function can change by switching modes or contexts. The top touch screen section can also contain multiple pages of actions and the rotary knobs change along with them, while the physical buttons are reserved for actions that should always be available. The control surface in its default configuration for Adobe Premiere Pro is presented in figure 5.1.

Figure 5.1. *Loupedeck CT*

The six small rotary knobs have a prominent tactile bump as they are rotated, giving feedback when an input is sent to the software. The centrepiece of the control surface is the rotary knob/touch screen control. The knob works as a jogging control, although fast jogging is not possible and has to be done with keypresses, and the surface of the knob is not grippy so continuous rotation is difficult. The touchscreen is mostly useful for colour grading where it controls the colour wheels for shadows, midtones and highlights.

The control surface provides a lot of feedback about its own state, but it is almost entirely visual and some of the buttons are on a touch screen, so removing gaze from the software is sometimes necessary. The biggest downfall is ergonomics in sustained use. Because the centre of the control surface is taken up by a sensitive rotary and touch control, the user has to "hover" their hand above it instead of resting and has to reach over for the touch panel controls.

Another issue is in the consistency of actions and error prevention: Some actions will activate a different panel in the software, disabling other actions completely until the timeline panel is reactivated by clicking on it. There is no clear indication of why this happens, rather it is something that the user has to be aware of and check for if an interaction on the controller invokes no action in the software. The checklist for the Loupedeck CT is presented in table 5.23.

Table 5.23. *Loupedeck CT heuristic evaluation*

Heuristic and checklist	Fulfilled
CS1: Presentation of feedback and system status	
If the control surface has multiple states, it communicates the state to the user	X
The control surface communicates information about the NLE state to the user	
CS2: Representative interaction	
Physical controls match the virtual controls of the NLE user interface	X
Controls are selected so that they are familiar to the user in the context of the action (e.g. slider for volume)	X
Interactions relate to the physical world where possible	

Continued on the next page

Table 5.23 – *Continued from previous page*

Heuristic and checklist	Fulfilled
CS3: Learnability	
The control surface comes with clear documentation for getting started	X
From the moment of unboxing, the first meaningful use of the control surface happens within an hour	X
A clear path for developing skills and improving expertise with the control surface is provided	X
CS4: Memorability and recognition	
The user interface provides unobtrusive hints of the users position on it, allowing the user to recognize the actions they are searching	X
The control surface allows users to memorize certain actions or features and execute them consistently	X
CS5: Predictability and consistency	
Each action on the control surface always has the same result	
If a control's role is dependent on the NLE or the state of the control surface, the state and outcome is clearly communicated	X
CS6: Flexibility and efficiency of use	
The control surface has basic functionalities easily available	X
The control surface offers shortcuts for advanced users, but their use is not required	X
CS7: Individualization and user freedom	
The actions of the control surface can be changed to user preference	X
Available actions that can be assigned to the control surface cover a broad range of approaches and workflows	X
CS8: Error prevention	
The control surface has indicators for what action a control will execute in a given context	X
The control surface only uses controls for specific actions that do not change	
The overall design prevents errors, such as unwanted tap or press gestures	
The control surface provides a way for error recovery	X

Continued on the next page

Table 5.23 – *Continued from previous page*

Heuristic and checklist	Fulfilled
CS9: Sustainable ergonomics	
The user can adopt different body positions while using the control surface	X
The body position required for using the device is resting and does not require constant muscle activation	
The control surface is accessible to people with tactile or motor disability	
CS10: Fit and finish	
The control surface feels robust and has a durable construction	X
The product looks professional	X
Maintainability has been taken into account in the design	X
CS11: Co-operation with other input devices	
The control surface is small enough to be used alongside a text input and pointer device	X
The control surface provides text input or pointer interaction	

5.2.2 Tunks Ergo33

Tunks Ergo33 is an open-source hardware design for a one-handed programmable USB keyboard (Kuitunen 2020). The keyboard uses QMK, a popular open-source firmware for mechanical keyboards that allows customizing any action (Humbert 2022). It is meant to mostly replace the keyboard while editing, but text input still needs to be done on a full-sized keyboard. The keyboard is presented in figure 5.2.

Figure 5.2. *Tunks Ergo33*

The keyboard is designed to be used with minimal hand movements: the keys are in a columnar layout, meaning straight columns of keys are offset vertically to fit the different lengths of fingers (Deskthority 2022). The hand is meant to be placed in the "WASD" area, commonly used in video games, and the there is one extra column to the left and two extra columns to the right of this position for the fingers to reach. The thumb has its own cluster with three keys. In the default mode the keyboard acts like the left half of a regular, QWERTY layout keyboard with an extra column to the left, but by pressing the rotary encoder the mode is changed to video editing and the keys are remapped to cover the most common keyboard shortcuts, marked on the sides of the keys. The status is indicated with the front-facing LEDs. The keyboard is fully programmable and allows the creation of multiple layers for different commands and additional status lights, so there is a lot of customization options available.

The keyboard is preloaded with a keymap with normal and editing mode, but any changes to the keymap have to be made by creating a new version of the firmware with an online configurator tool and flashing the firmware onto the device. This is not user-friendly, but many other QMK keyboards use the VIA software for easier configuration and for a commercial product this would solve the issue. Ergo33 provides good tactile feedback about the position of the hand but feedback about the control surface's state is only provided by the multi-coloured lights. This limits the amount of different modes or layers a user can realistically put on the keyboard, as the user has to remember the layout of each layer or have it marked on the keycaps. The rotary encoder is also an outlier to the ergonomics of the control surface. Using it would require the user to remove their hand from the home position, so this results in the encoder being left unused. The layers and the rotary knob are the only real differences between using the Ergo33 and a normal keyboard, as all modern editing software allows the user to customize the keyboard shortcuts. Portability to different computers could be another advantage, as the Ergo33 will work with the default shortcuts of the software and all of the remapping is done inside the keypad. The build quality of the hardware does not match mass-produced products because tooling for an outer shell is expensive and time-consuming, but the device still feels relatively robust. The checklist for the Ergo33 is presented in table 5.24.

Table 5.24. *Tunks Ergo33 heuristic evaluation*

Heuristic and checklist	**Fulfilled**
CS1: Presentation of feedback and system status	
If the control surface has multiple states, it communicates the state to the user	X
The control surface communicates information about the NLE state to the user	
CS2: Representative interaction	
Physical controls match the virtual controls of the NLE user interface	
Controls are selected so that they are familiar to the user in the context of the action (e.g. slider for volume)	
Interactions relate to the physical world where possible	
CS3: Learnability	
The control surface comes with clear documentation for getting started	
From the moment of unboxing, the first meaningful use of the control surface happens within an hour	X
A clear path for developing skills and improving expertise with the control surface is provided	X
CS4: Memorability and recognition	
The user interface provides unobtrusive hints of the users position on it, allowing the user to recognize the actions they are searching	X
The control surface allows users to memorize certain actions or features and execute them consistently	X
CS5: Predictability and consistency	
Each action on the control surface always has the same result	
If a control's role is dependent on the NLE or the state of the control surface, the state and outcome is clearly communicated	

Continued on the next page

Table 5.24 – *Continued from previous page*

Heuristic and checklist	Fulfilled
CS6: Flexibility and efficiency of use	
The control surface has basic functionalities easily available	X
The control surface offers shortcuts for advanced users, but their use is not required	X
CS7: Individualization and user freedom	
The actions of the control surface can be changed to user preference	X
Available actions that can be assigned to the control surface cover a broad range of approaches and workflows	X
CS8: Error prevention	
The control surface has indicators for what action a control will execute in a given context	X
The control surface only uses controls for specific actions that do not change	
The overall design prevents errors, such as unwanted tap or press gestures	X
The control surface provides a way for error recovery	X
CS9: Sustainable ergonomics	
The user can adopt different body positions while using the control surface	X
The body position required for using the device is resting and does not require constant muscle activation	X
The control surface is accessible to people with tactile or motor disability	
CS10: Fit and finish	
The control surface feels robust and has a durable construction	
The product looks professional	
Maintainability has been taken into account in the design	X
CS11: Co-operation with other input devices	
The control surface is small enough to be used alongside a text input and pointer device	X
The control surface provides text input or pointer interaction	

5.2.3 Wacom Intuos Pro M

Wacom Intuos Pro M is a drawing tablet with a fairly large drawing area and some additional buttons. Drawing tablets in general have garnered popularity and attention in the creative digital arts, even to the point where satiric videos have been made about them (Vae 2011), since early 2010s. It acts as a pointer device so depending on the setup it can fully replace a mouse, and in all use cases the user will be using either a mouse or the drawing tablet, not both at the same time. The drawing area can be configured to cover multi-display or a single display only. The side of the drawing area also contains programmable buttons that display their actions on an on-screen display if the user hovers their fingers over the buttons. The drawing tablet is presented in figure 5.3.

Figure 5.3. *Wacom Intuos Pro M*

The drawing tablet's main interaction is using the pen as a pointer device, and unlike graphic design or digital painting, in video editing this interaction does not change the efficiency of the user significantly. The main advantage of the control surface compared to mouse is the different ergonomics, which for some users will be a better fit for sustained use. Besides the different interaction the programmable macro buttons are once again proven useful, as they can be configured differently for each software. The biggest drawback of the pen interaction is accuracy: because the usable area of the drawing tablet is much smaller than the screen area it is mapped to, even a small unintended movement with the pen, for example while bringing the tip down from hover to do a click, will move the cursor. Double-click with the pen proved to be almost impossible during the evaluation session, and while accuracy would most likely increase through continued use, this caused frustration in the beginning. The situation would be different if the display and the drawing surface were one. Hollywood editor Alan Bell is known for using a device like this (Elwyn 2014) and one of the participants of the contextual inquiry expressed their interest in this kind of device instead of a separate drawing tablet and a monitor. The checklist for the Intuos Pro M is presented in table 5.25.

Table 5.25. *Wacom Intuos Pro M heuristic evaluation*

Heuristic and checklist	Fulfilled
CS1: Presentation of feedback and system status	
If the control surface has multiple states, it communicates the state to the user	X
The control surface communicates information about the NLE state to the user	
CS2: Representative interaction	
Physical controls match the virtual controls of the NLE user interface	
Controls are selected so that they are familiar to the user in the context of the action (e.g. slider for volume)	
Interactions relate to the physical world where possible	
CS3: Learnability	
The control surface comes with clear documentation for getting started	X
From the moment of unboxing, the first meaningful use of the control surface happens within an hour	X
A clear path for developing skills and improving expertise with the control surface is provided	X
CS4: Memorability and recognition	
The user interface provides unobtrusive hints of the users position on it, allowing the user to recognize the actions they are searching	X
The control surface allows users to memorize certain actions or features and execute them consistently	X
CS5: Predictability and consistency	
Each action on the control surface always has the same result	X
If a control's role is dependent on the NLE or the state of the control surface, the state and outcome is clearly communicated	

Continued on the next page

Table 5.25 – *Continued from previous page*

Heuristic and checklist	Fulfilled
CS6: Flexibility and efficiency of use	
The control surface has basic functionalities easily available	X
The control surface offers shortcuts for advanced users, but their use is not required	X
CS7: Individualization and user freedom	
The actions of the control surface can be changed to user preference	X
Available actions that can be assigned to the control surface cover a broad range of approaches and workflows	X
CS8: Error prevention	
The control surface has indicators for what action a control will execute in a given context	X
The control surface only uses controls for specific actions that do not change	
The overall design prevents errors, such as unwanted tap or press gestures	
The control surface provides a way for error recovery	X
CS9: Sustainable ergonomics	
The user can adopt different body positions while using the control surface	X
The body position required for using the device is resting and does not require constant muscle activation	X
The control surface is accessible to people with tactile or motor disability	
CS10: Fit and finish	
The control surface feels robust and has a durable construction	X
The product looks professional	X
Maintainability has been taken into account in the design	
CS11: Co-operation with other input devices	
The control surface is small enough to be used alongside a text input and pointer device	
The control surface provides text input or pointer interaction	X

5.3 Summary

The heuristics used for evaluation were refined during the definition stage, but the full process would require an additional validation and refinement stage to be iterated until the heuristics evaluate the application domain sufficiently. The heuristics created during this process evaluate video editing control surfaces well, but the checklists would have likely benefited from further refinement. Some of the checklist items are ambiguous and there should be a distinction between standalone items and items that are a part of a group, where some of the items are mutually exclusive. This is partially an issue with the domain of the evaluation: the design and purpose of different control surfaces vary greatly, so it is difficult to come up with a set of heuristics that are specific enough to bring out the differences between devices and potential usability issues, and at the same time be generic enough to evaluate devices meant for completely different use cases. Additionally, the granularity of the checklist is not enough to make distinctions between how well two different user interfaces fulfil it. This is not necessarily the purpose of the checklist and these differences can be described in detail in addition to the checklist.

The methodology for creating heuristics by Quiñones et al. works well at steering the process and forcing the researchers to justify their decisions with the products of the process. As long as the information gathered in the first stages is sufficient, the descriptive and correlational stages emphasize the significant features of the domain clearly, and it is easy to decide which features the heuristics should cover based on the results of these stages. The heuristics defined in section 5.1.5 could also be used, with some modifications, to evaluate systems containing physical user interfaces in different application domains.

6. DESIGNING AND EVALUATING A NEW CONTROL SURFACE

For the design and evaluation phase of the research, the original plan was to first create the concept for a new control surface, evaluate it through a low-fidelity prototype, then refine the design and produce a physical, high-fidelity prototype that could be used for user testing. However, based on previous results it was clear that in order to properly evaluate the usability and user experience with a physical device, the whole system containing the device, its software and the integration with the editing software would all have to be prototyped to a high degree of finish.

In the context of video editing, especially with expert users, the final polish of the whole system matters a lot to the efficiency of use and all facets of the system would need to be refined through an intense product design cycle. This level of polish was not achievable in the scope of this research and evaluating a physical prototype without this refinement would not bring any additional value or results to the research. Because of these findings, the physical prototype and the subsequent evaluation was excluded from the research and the final phase focuses on exploring novel interaction techniques for video editing through the use of a low-fidelity prototype.

6.1 Prototype and Evaluation

Based on the previous results there are two distinct categories of actions in editing software: discrete actions, such as splitting or deleting a clip and play/pause, and continuous actions, such as navigating the timeline, scrubbing or jogging, adjusting clip volume, and dragging a clip or an edit on the timeline. Discrete actions are difficult to vision new interactions for, as a simple button input is efficient and most of the potential improvement is in the visibility of available actions and feedback. The Elgato Stream deck accomplishes this with displays behind every button (Elgato 2022) while the Loupedeck CT uses a touch screen for some actions, and in this case it is hard to avoid using visual feedback. Technically even the continuous actions have a limited resolution, but they can be done on an analogue control that allows a continuous range of adjustment. The control should provide enough precision to be useful, but at the same time not force the user to excessive motion.

The evaluated concept should incorporate both discrete and continuous controls and the most interesting possibilities lie within the continuous controls. Because of this the concept is based on a programmable keyboard with a selection of continuous controls available, namely a joystick, a rotary knob and mid-air gestures. This combination will give participants of the evaluation a familiar baseline, some controls they are likely to be familiar with but possibly not in this context, and some controls they likely have not tried themselves. The prototype would give examples on what tasks a certain control might be useful for, but as stated in chapter 4, customizability is an important factor for editors, so the prototype should demonstrate this as well.

The prototype was created by visually modifying an existing control surface, the Tunks Ergo33, to represent new interactions and then recording short videos (6-20 seconds) of the interactions with a simultaneous screen recording of Adobe Premiere Pro to show the effects of each interaction. These videos were used in an online survey, where the participants would watch each video and then evaluate the interaction on attractiveness, usefulness and willingness to spend both money and time to obtain and learn the control. The evaluations were done on a 5-point Likert scale from strongly disagree to strongly agree, and an optional open answer field. The survey also included questions about the participants' backgrounds in video editing, experience with editing control surfaces, and general opinions about adopting new control surfaces into their workflow. The subjects of the videos are presented in table 6.1, but videos 4 and 5 were evaluated together as they demonstrated two different use cases for the same interaction. The structure of the survey is presented in appendix C and example screenshots of the prototype videos are presented in appendix D.

Table 6.1. *Low-fidelity prototype interactions*

Video	Interaction technique	Interaction category	Task
1	Key press with changeable layers	Discrete	Jogging (playback forwards and backwards), frame by frame navigation
2	Joystick	Continuous	Timeline navigation: horizontal scroll and zoom
3	Rotary wheel	Continuous	Slipping a clip
4	Mid-air gestures	Continuous	Clip transform: rotation and position
5	Mid-air gestures	Continuous	Editing keyframes: clip volume

6.2 Results

40 participants answered the survey overall and of them 15 were amateurs and 25 professional editors. The participants rated their own level of experience and most of them considered themselves intermediate or advanced. Significantly more professionals considered themselves advanced or expert, as shown in figure 6.1.

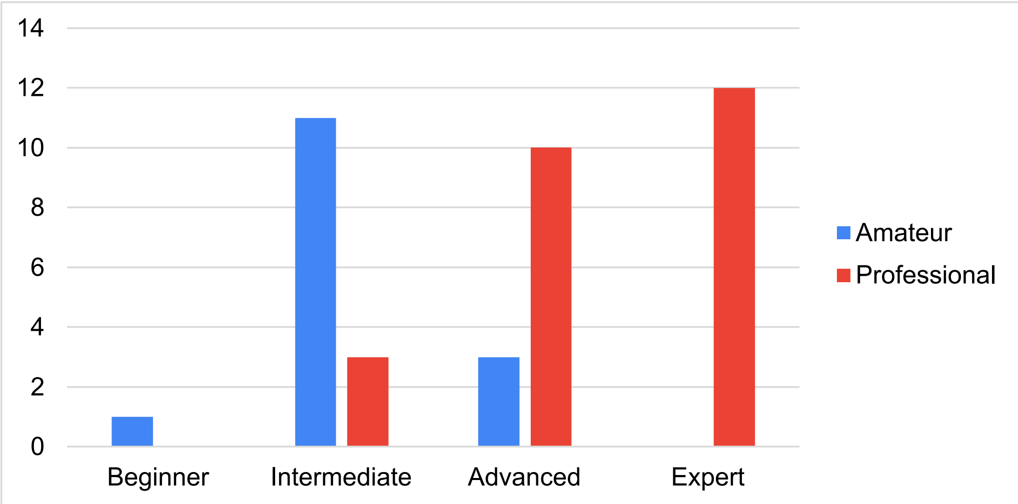


Figure 6.1. *Survey participants' level of experience*

Adobe Premiere Pro was the most common software among the participants, but it is worth noting that some editors use multiple software, up to 4 for one participant. Figure 6.2 shows the proportion of participants using a particular software, but the options are not exclusive. 11 of the 40 participants use an additional controller besides keyboard and mouse and the answers can be divided into three categories: a macro keypad, a keypad with a jog control like the Loupedeck CT, or a drawing tablet. The usage of additional control surfaces was evenly divided between amateur and professional editors when taking into account the disparity in the number of answers from each group.

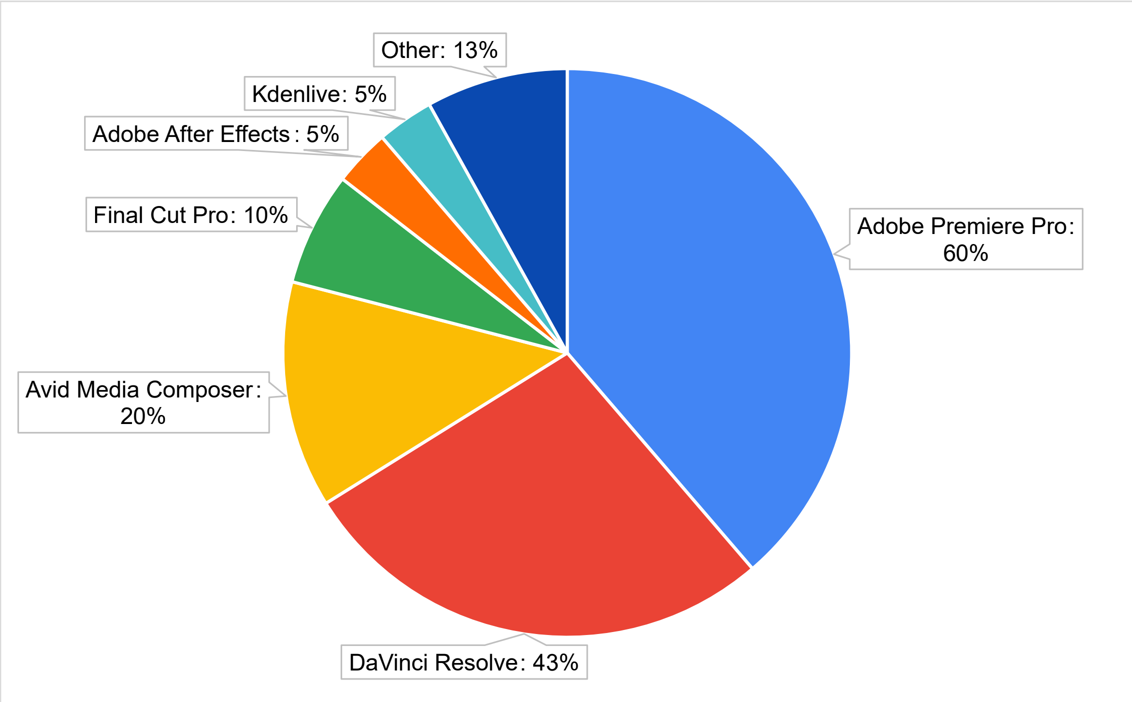


Figure 6.2. *Software used for editing*

In general, professional editors were more reluctant to adopting any of the prototyped interactions. This was reflected in all questions on the scale: interest in the control, perceived usefulness and willingness to spend time or money on the control. However, some participants had misunderstood the point of the prototype which was to evaluate the interaction method, not the specific action or task that was used as an example. All of the graphs are presented as percentage of the participant group (amateur or professional) answering a specific option on the Likert scale.

The switchable layers on a keyboard gained a lot of interest from amateurs as seen in figure 6.3, with the majority agreeing that this would be both interesting and useful. Some 50% the professionals were not interested, either disagreeing or strongly disagreeing with all claims. The open answers for this interaction were mostly positive, mentioning how this would be useful when the user is running out of available keyboard shortcuts, but some participants preferred using modifier keys like Control, Alt and Shift, or having dedicated controls for every action without any modifiers.

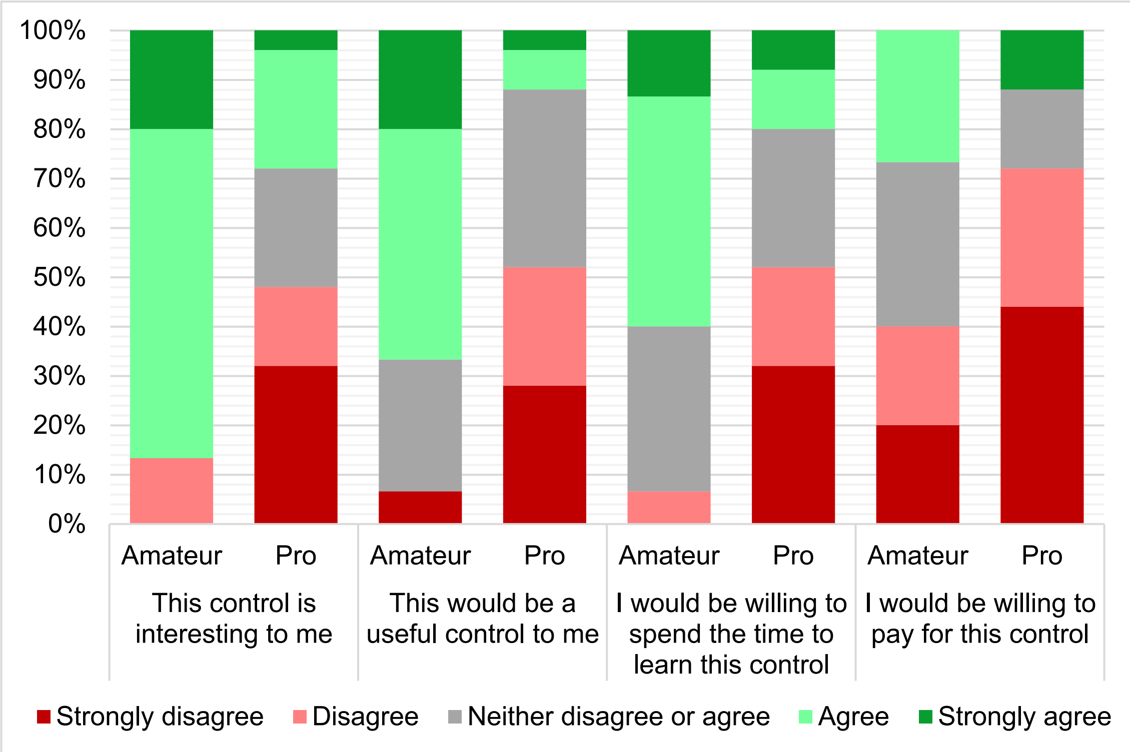


Figure 6.3. *Keyboard with switchable layers*

The joystick was less popular than the switchable layers among the participants, with some 60% of the amateurs indicating some level of interest, but in every other category less than 50% of both user groups agreed on the usefulness or willingness to spend time or money on it. The results are presented in figure 6.4. The open answers all indicated that this would be useful for timeline navigation, meaning scrolling and zooming in both axes and moving the playhead. Some computer mice offer a horizontal scroll wheel which some participants used for the same purpose.

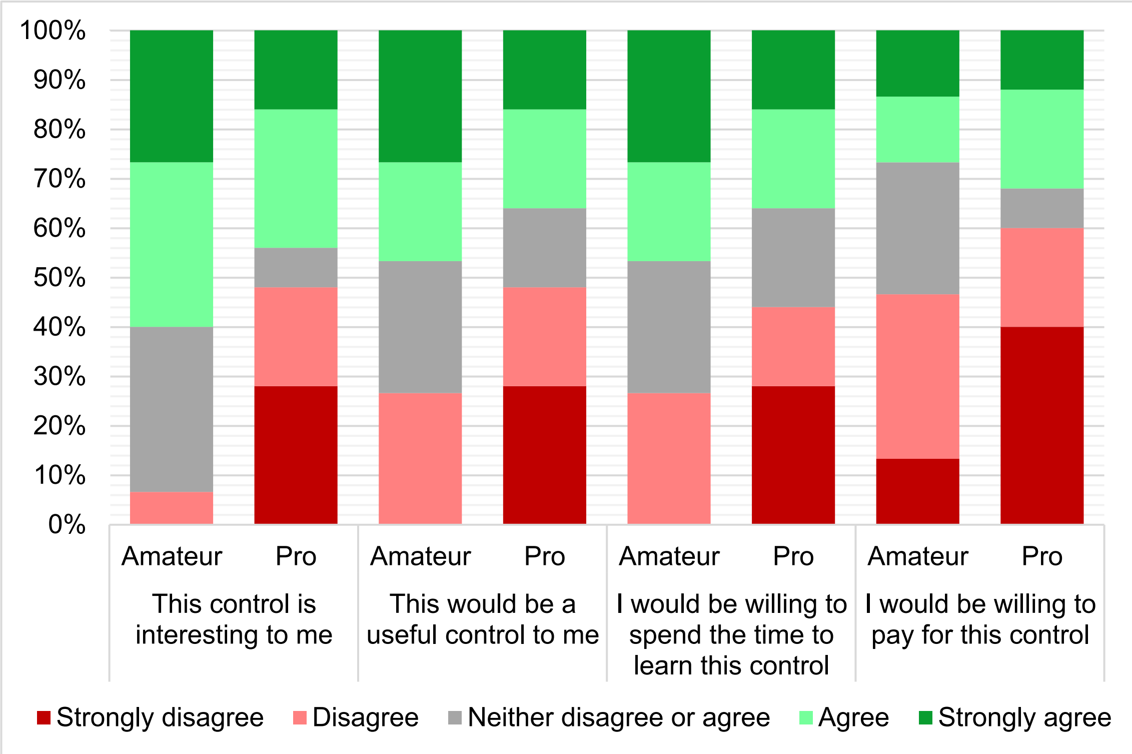


Figure 6.4. *Joystick*

The rotary wheel was by far the most popular of the prototyped interactions, although again the professional editors were mostly not interested as seen in figure 6.5. The open answers were divided between the control making some tasks easier, mouse and keyboard being able to do the same tasks, and concerns about the tactility of the wheel itself. The importance of grip on a control like this was noted in section 5.2.1 and the survey results seem to back this point.

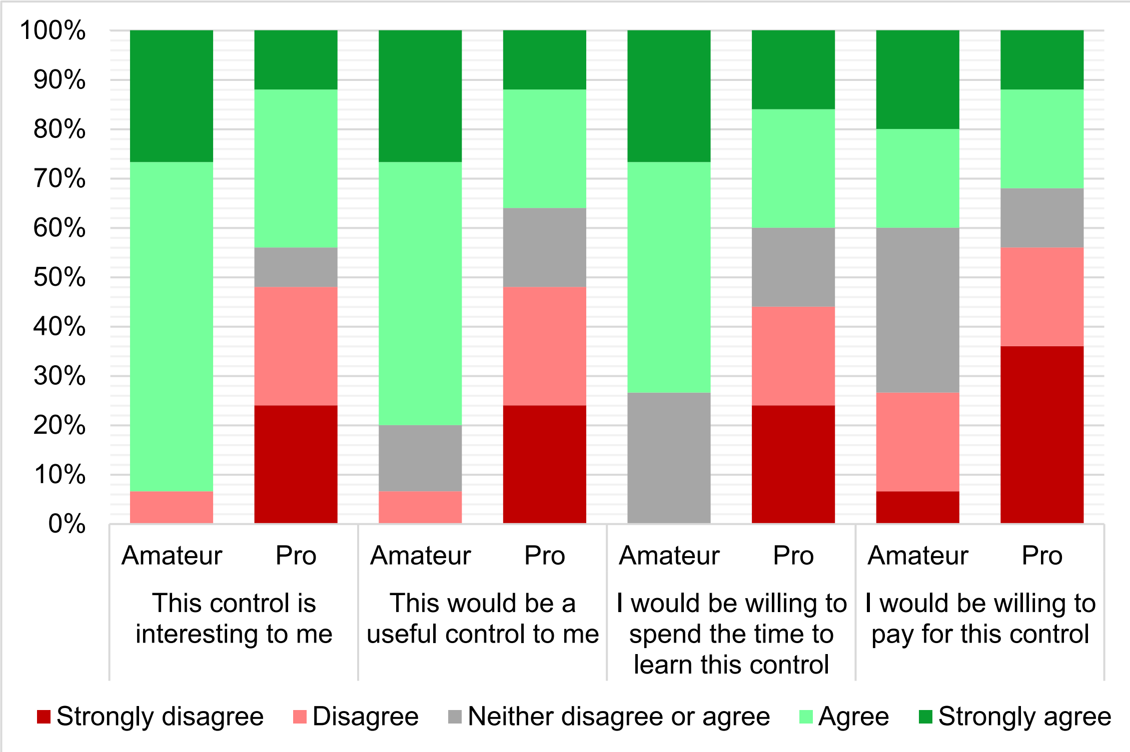


Figure 6.5. *Rotary wheel*

The mid-air gestures received by far the most negative reaction across both groups, as seen in figure 6.6. The complete lack of tactile feedback and the resulting loss of accuracy seemed to be the biggest issue for many participants. Others considered the interaction method interesting but not necessarily useful. One participant noted experimenting with various mid-air tracking systems like the Leap Motion, but that frustration with the lack of support and arm fatigue were the primary reasons for giving up on them. Interestingly, one professional editor was exceptionally excited about the prospect of using mid-air gestures where most participants considered the idea not interesting.

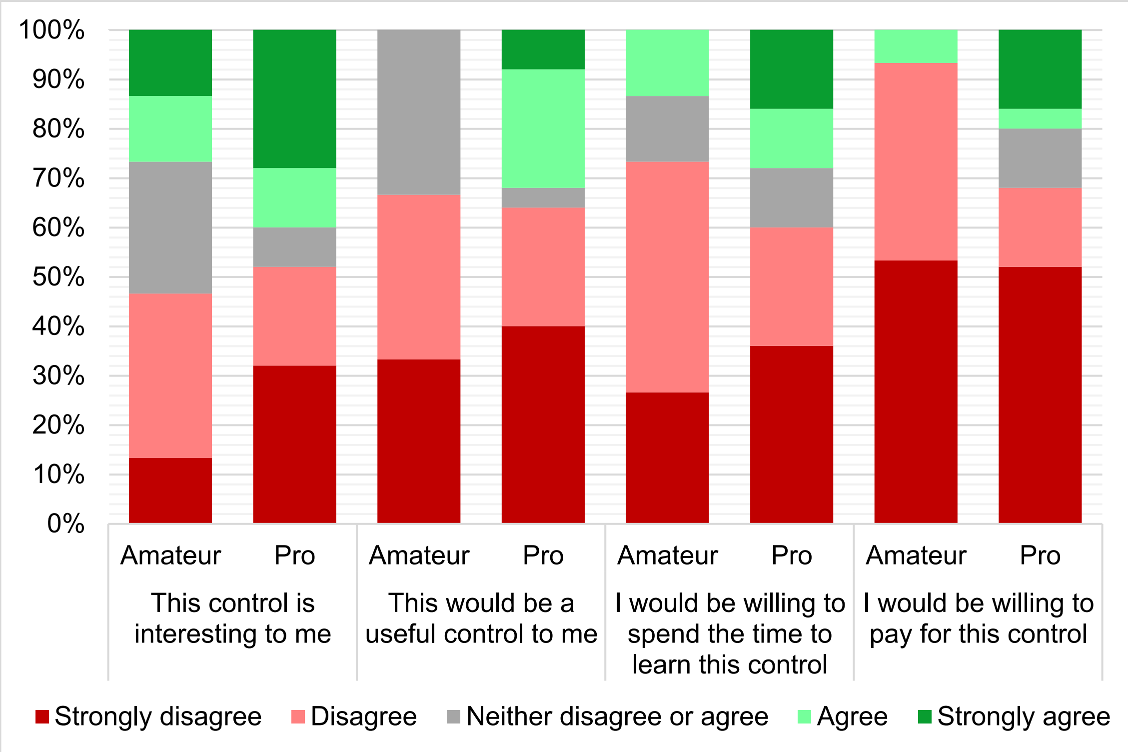


Figure 6.6. *Mid-air gestures*

In the final part of the survey, the participants were asked about their general opinions on video editing and editing-specific control surfaces. Figures 6.7 and 6.8 show that amateur editors were more interested in specialized control surfaces and willing to spend time learning a new product, but professional editors were still more willing to spend money on one than amateurs. The majority of both user groups were more interested in using familiar interaction methods over novel ones. The majority of both groups also agreed that keyboard and mouse are good enough for video editing, despite the interest in specialized control surfaces. It is not surprising that since a larger proportion of professional editors considered their skills advanced or expert level, fewer professionals would be capable or willing to improve their editing workflow as seen in figures 6.8 and 6.9. The distinction in the results is that some of the participants capable of improving their workflow are not willing to do it, but interestingly, neither user group prioritized editing over improving the workflow strongly and opinions seemed to be divided. The last question of the usefulness of specialized control surfaces compared to experience level seemed to divide the user groups the most, with a significantly larger part of professionals strongly disagreeing with the claim that the control surfaces would be more useful to professionals, while a larger part of amateurs agreed with the claim.

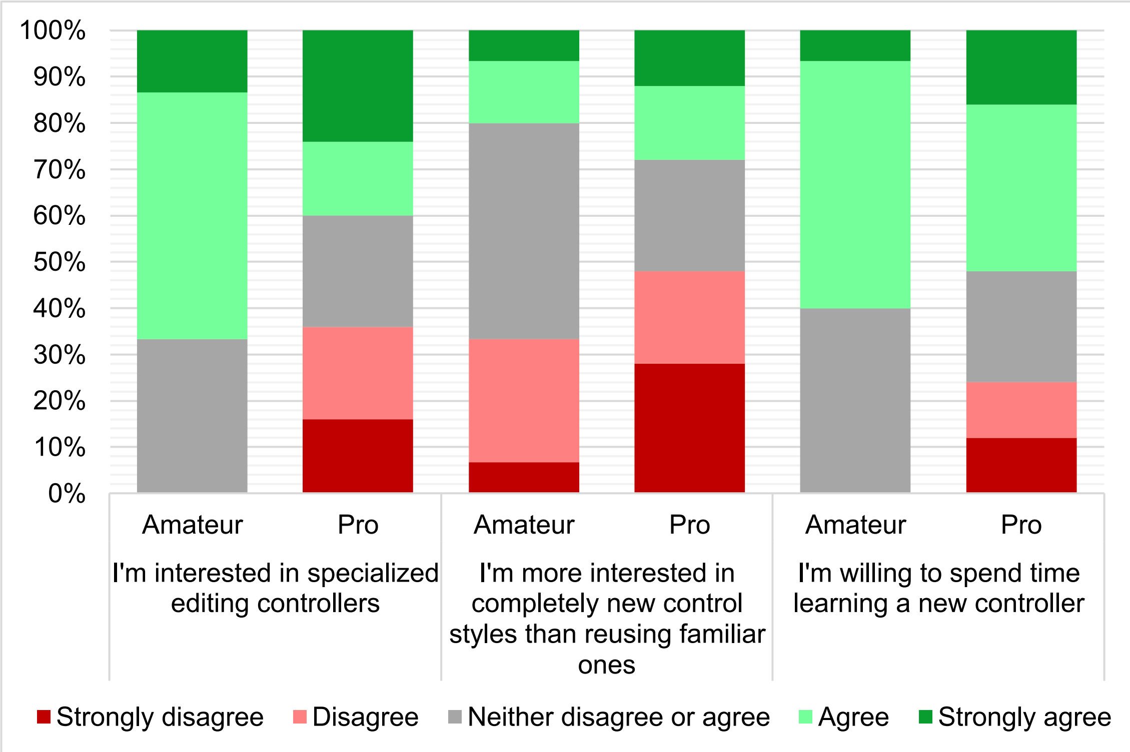


Figure 6.7. *General opinions part 1*

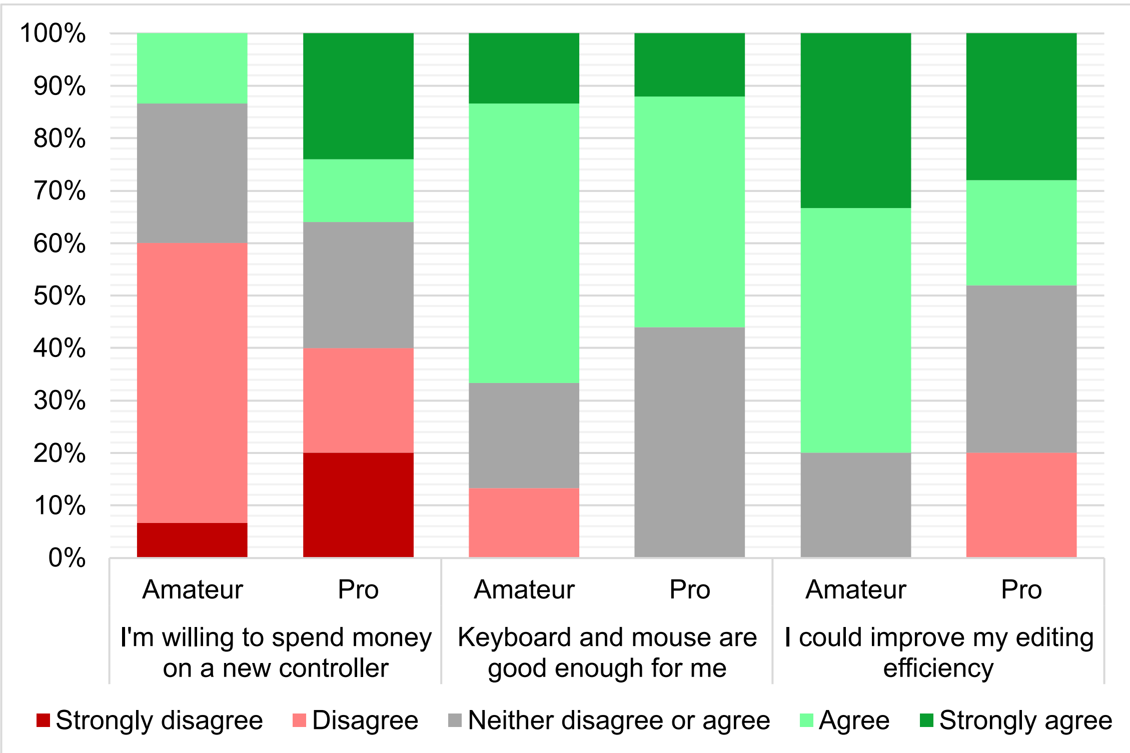


Figure 6.8. *General opinions part 2*

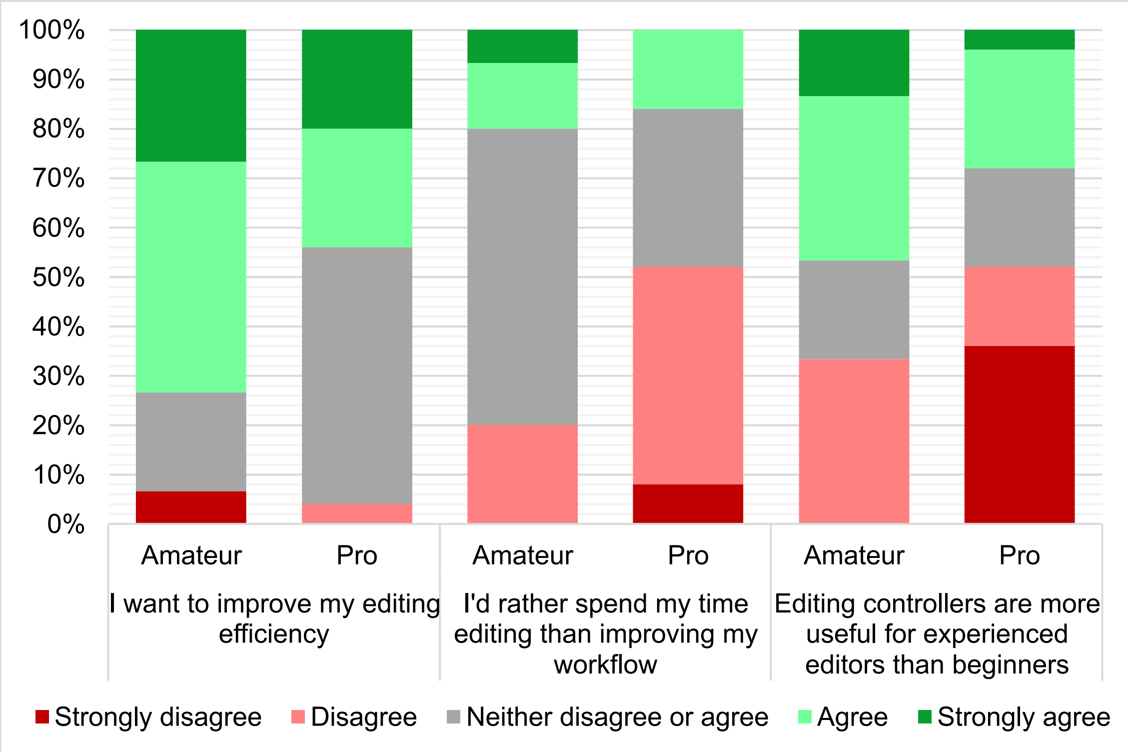


Figure 6.9. *General opinions part 3*

Multiple open answers at the end of the survey showed concern about support and maintainability for the hardware and software, and integration into the editing software. Ergonomics are also brought up, one participant specifically naming carpal tunnel syndrome as a potential issue with some interaction methods. There is also a good point brought up about the workflow: editing is about more than just the timeline and the picture, as there are also many audio tracks and audio or video plugins that are a part of the process, and a good controller would also help with those use cases.

One idea is brought up about voice commands as a modality, for example for tasks usually done with keyboard shortcuts. This research was based on the hypothesis that audio as a modality is undesirable and audio is better left for monitoring and feedback. This answer indicates that there might be some interest in audio as an input method, and it would also increase the accessibility of video editing to people with reduced tactile function, possibly combined with gaze tracking to replace the pointing device.

7. DISCUSSION

In this chapter the main findings of the research are discussed and based on the findings and the products of previous stages, the design guidelines are formed. Finally, the limitations of the research are discussed.

7.1 Main Findings

The first research question of the thesis was what is important to the user in a video editing control surface. The most important factors are flexibility, learnability and findability, and ergonomic sustainability. There are as many combinations of software and workflows as there are editors, so customizability is crucial to get existing users to adopt a new control surface. Supporting the users' current workflows enough that the control surface is appealing to them is also important. For novice users without many habits, it might be possible to introduce a new way of working along with a new control surface, but inexperienced users are not likely to buy a specialized device for video editing and instead start with the mouse and keyboard. None of the previous research presented in section 2.4 focused on existing editing software, so the user needs found in this research are very different compared to research on novel video editing techniques.

There are three different aspects to learnability and findability: the initial learnability, memorability and the resulting final plateau of efficiency, and findability of the peripheral actions. Video editing is a task that requires in-depth knowledge, but the control surface should not increase this initial learning curve. It should be easy to adopt a new control surface and use it meaningfully, while still offering a path for further learning and personal development. One of the most important features of any control surface is to allow the user to concentrate their most used actions into controls that are easily reachable and memorizable, and allow the user to find the less common actions easily through various cues. The final plateau of efficiency is reached when the user has developed sufficient memory for the control surface, and this ceiling should be raised as high as possible without complicating the learning process. Professional and amateur editors both may be willing to learn new control surfaces and interaction methods, but novel controls are not necessarily interesting and the value proposition, in essence the increase in efficiency in exchange for time spent learning, needs to be clear even before use. This finding correlates with previous research from Diogo Cabral and Nuno Correia, where an unfamiliar user interface was considered attractive and interesting by many users, but not necessarily efficient (Cabral and Correia 2017).

The second research question was what is the right degree of specificity for a control surface. The answer is two-fold: an individual control on the control surface can be highly specialized for certain action, but as a whole the control surface should support different tasks and phases of editing. The editor will also be doing other tasks outside the editing software as a part of their work and these actions will likely need text input and pointing interaction. Either the control surface facilitates these input modalities or works in conjunction with other input devices to fulfil the need. The degree is specificity is a question of balance and varies for each control surface. Generally speaking, the more specific a control surface is, the clearer its increase in efficiency needs to be in that task and the more common that task needs to be for the user.

The final research question was what design considerations does the task of video editing cause. One of the most important considerations is the nature of the underlying action: in the domain of video editing, all actions are divisible into discrete or continuous actions and this sets limits on what interaction techniques make sense for each action. This distinction of action types in video editing was not found in the previous research presented in section 2.4, likely because much of the research was focused on completely new ways of video editing and less experienced users. A change in the software UI can affect what controls are suitable for a specific task, but to change the action from one type to the other, the entire nature of the action would need to change. This means, to some extent, redesigning the principles on which the software operates. The design guidelines for video editing control surfaces also answer this research question and they are discussed in section 7.2.

One important finding that is not directly related to any of the research questions is the overall need for a specialized control surface for video editing. Based on the contextual inquiry and the prototype evaluation, keyboard and mouse are sufficient for most editors and while keyboard shortcuts require a lot of memorization, the end result is that a frequent user will become so efficient that the bottleneck is the creative problem-solving, not the execution of individual tasks. This is mostly due to the fact that a keyboard is already an efficient control surface for discrete actions and the mouse and its scroll wheel can be used for the continuous actions.

7.2 Forming the Design Guidelines

The design guidelines were formed to answer the third research question of design considerations more comprehensively. The guidelines were formed by reviewing the outcomes of the research phases and utilizing the products of the research, specifically the personas created in section 4.5, the heuristics defined in section 5.1.5, and the prototype evaluation results from section 6.2. The guidelines share a lot of similarities with the heuristics but focus on guiding the design process instead of evaluation.

1. **Define the product's focus and the interoperability with other products.** It is difficult to cover every possible task with a single control surface, so it is best to select which parts of the workflow to support. The focus still needs to be broad enough to appear useful to the users. Additionally, the editor will need the ability to do tasks outside the editing software, so your product should facilitate these or allow the user to use other devices for these interactions.
2. **Design for flexibility.** Allow the users to customize the control surface and accommodate their existing workflows. The customizability may be in the software only, but some users will also be interested in customizing the hardware by for example reorganizing the controls.
3. **Use appropriate interaction types for your planned actions.** Use discrete or continuous interactions for actions where they make sense, match the physical control to the editing software where possible, and consider the users' familiarity and expectations with different control types.
4. **Provide a quick way to start and a clear path for learning.** The first use of the product should be effortless and lead to meaningful results quickly to lower the threshold for adoption. The primary value proposition of the product is the long-term increase in efficiency, so a path for developing beyond the initial use should also be provided.
5. **Build memorability and findability into the user interface.** Allow the user to memorize and non-visually recognize their most used actions, and provide a convenient way to find the less common actions with cues. The user should focus on the output of the editing software instead of the usage of the control surface.
6. **Design for sustainable long-term use.** Active amateur and professional editors spend a lot of time editing for multiple days in a row, so the use of the control surface should strain the user as little as possible, and not aggravate any issues caused by repetitive strain.
7. **Ensure the sum of the physical device, the companion software and the integration with editing software works.** The product consist of these three parts in equal measure, and any of the parts failing will lead to the product as a whole failing.
8. **Design support as a part of the product.** Users are concerned about the long-term viability of their investment of time and money. The hardware should be robust and repairable, and the software and integration should be actively supported. Products that lack certainty will discourage users from making the investment.

7.3 Limitations

This research was done in the context of current desktop-based editing software. The user interface and available controls of the software, as well as the physical desktop environment, limit the features of the control surface to what makes sense in the current context. While the research does not address any particular tasks in detail, it is possible for the software and the environment of use to change dramatically in the future and this would also change the needs of the users. Handheld smart devices are already capable of video editing and it is possible for them to take over simple editing tasks using footage recorded on the device, but more complex editing projects and editing footage from dedicated cameras will most likely remain on computers due to the processing power and storage capacity requirements. Another related limitation is the users' experience with current editing software. Especially for experienced users it can be difficult to imagine another way of working, but if a new control surface is being designed for the current market, these are the potential customers the product has to convince.

The lack of members in the research team is a major limitation. This results in the lack of multidisciplinary collaboration and multiple points of view for large parts of the research, even though several people contributed to the research in the form of discussions and advice during the process. It would have also been beneficial to include some of the stakeholder groups in the research, especially the developers of the editing software on the market. The software integration of a control surface is instrumental to its success and understanding what goes into a successful integration would likely lead to more specific related guidelines for the design process.

Video editing is also a topic of deep knowledge and a variety of user needs. A short contextual inquiry session may not give the researcher sufficient context for the design process, unless they have prior experience of video editing themselves. In this case prior experience was present which helped with this part of the research, but it may also cause a bias in the formulation of questions and the analysis of the results. A deeper understanding of the domain and the end users can be built through a longer, iterative product design process, but this was not possible due to the scope of the research.

8. CONCLUSION

The first motion pictures were screened to audiences across the world in the end of the 19th century (Knight 1957, pp. 18–19) and many technical developments were made in the following decades up until the 1950s. At that point, the medium of film had already reached much of the potential it would ultimately have (Knight 1957, pp. 288–298). The process of editing, however, would remain the same difficult and tedious task for a large part of the century. The introduction of the videotape and non-linear editing revolutionized the work of the editor and digital media and tools took the capability of a single person further than ever before, but this period in the history of film and video editing has only existed for a short amount of time. Technology now is advancing faster than ever before and video editing has been democratized: no longer do you need an expensive workstation or complicated software for it, although complex video editing for large projects will likely still remain the domain of expert users and desktop software.

Video editing by nature is full of discrete actions: the media formats we edit are limited in their spatial and temporal resolutions, and editors often end up working right at the limit of that resolution. The results of this research demonstrated that in the context of modern NLEs, there is no benefit in trying to map these discrete actions to a continuous analogue interaction, as it will only decrease accuracy. The continuous actions show potential for new interaction design along with voice control as a new input modality, but the level of prototyping required to validate new concepts in this context make it difficult to determine which ideas could become the next progression in video editing. In addition, there might not be much need for editing-specific control surfaces as modern editing software is designed to work well with a keyboard and mouse. The results showed that while some tasks would benefit from more specialized controls, they might not justify getting a separate control surface. Either the increase in efficiency might not be enough, or the task is not common enough for the user.

The user's workflow is a result of their responsibilities in a given project, their previous experience, and their preferences. Video editing also happens in a broad range of environments, from soundproofed editing suites with big screens to a laptop and a pair of headphones on the go. The design process for a control surface has to define the primary user group and find out their needs and context of use, because any design is a compromise between the priorities of different user groups and stakeholders.

REFERENCES

Avid Technology (1989). Avid Technology's New Digital Non-Linear Editing System Combines Film and Video Editing Techniques. In: Business Wire.

Budiu, Raluca (2014). Memory Recognition and Recall in User Interfaces. Nielsen Norman Group. Available (accessed on 24 February 2022): https://www.nngroup.com/articles/recognition-and-recall/.

Cabral, Diogo and Nuno Correia (2017). Video editing with pen-based technology. In: Multimedia tools and applications Vol. 76(5), pp. 6889–6914. DOI: 10.1007/s11042-016-3329-y.

Dancyger, Ken (2013). The technique of film and video editing: history, theory, and practice. 5th ed. Focal Press. ISBN: 9780240813974.

Deskthority (2022). Staggering. Available (accessed on 9 March 2022): https://deskthority.net/wiki/Staggering.

Dirk, Tim (2022). Film Terms Glossary. Available (accessed on 29 January 2022): https://www.filmsite.org/filmterms.html.

Doherty, Edward (2020). Editing FCPx with an Xbox controller. Available (accessed on 15 February 2022): https://northernlightswedding.com/wedding-blog/2019/11/15/editing-with-an-xbox-controller, archived at https://web.archive.org/web/20220215095240/https://northernlightswedding.com/wedding-blog/2019/11/15/editing-with-an-xbox-controller on 15 February 2022.

Elgato (2022). Stream deck. Available (accessed on 25 March 2022): https://www.elgato.com/en/stream-deck.

Elwyn, Jonny (2014). Film Editing Keyboards, Mouse, Controllers and more. Available (accessed on 15 February 2022): https://jonnyelwyn.co.uk/film-video-editing-tools-for-editors/avid/editing-gear-keyboards-mouse-tablets-and-more/, archived at https://web.archive.org/web/20220215101233/https://jonnyelwyn.co.uk/film-video-editing-tools-for-editors/avid/editing-gear-keyboards-mouse-tablets-and-more/ on 15 February 2022.

ISO 9241-5:1998 (1998). Ergonomic requirements for office work with visual display terminals (VDTs) – Part 5: Workstation layout and postural requirements. International Organization for Standardization. Available: https://www.iso.org/standard/16877.html.

ISO 9241-400:2007 (2007). Ergonomics of human–system interaction – Part 400: Principles and requirements for physical input devices. International Organization for Standardization. Available: https://www.iso.org/standard/38896.html.

ISO 9241-11:2018 (2018). Ergonomics of human-system interaction – Part 11: Usability: Definitions and concepts. International Organization for Standardization. Available: https://www.iso.org/standard/63500.html.

ISO 9241-110:2020 (2020). Ergonomics of human-system interaction – Part 110: Interaction principles. International Organization for Standardization. Available: https://www.iso.org/standard/75258.html.

ISO 9241-112:2017 (2017). Ergonomics of human-system interaction – Part 112: Principles for the presentation of information. International Organization for Standardization. Available: https://www.iso.org/standard/64840.html.

ISO 9241-171:2008 (2008). Ergonomics of human-system interaction – Part 171: Guidance on software accessibility. International Organization for Standardization. Available: https://www.iso.org/standard/39080.html.

ISO 9241-210:2019 (2019). Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems. International Organization for Standardization. Available: https://www.iso.org/standard/77520.html.

Harley, Aurora (2015). Personas Make Users Memorable for Product Team Members. Nielsen Norman Group. Available (accessed on 1 April 2022): https://www.nngroup.com/articles/persona/.

Hassenzahl, Marc (2004). The Thing and I: Understanding the Relationship Between User and Product. In: Funology: From Usability to Enjoyment. Ed. by Mark A. Blythe, Kees Overbeeke, Andrew F. Monk and Peter C. Wright. Springer, pp. 31–42. ISBN: 9781402029660.

Hiebert, Ray E., Gerard J. Tortora and Carlton E. Spitzer (1968). The Voice of Government. Ed. by Ray E. Hiebert and Carlton E. Spitzer. Environmental Science and Technology. John Wiley & Sons, Inc. ISBN: 9780471367550.

Holtzblatt, Karen and Hugh Beyer (2016). Contextual Design: Design for Life. 2nd ed. Morgan Kaufmann. ISBN: 9780128008942.

Humbert, Jack (2022). QMK Firmware. Available (accessed on 9 March 2022): https://qmk.fm/.

Joyce, Alita (2019). How to Measure Learnability of a User Interface. Nielsen Norman Group. Available (accessed on 24 February 2022): https://www.nngroup.com/articles/measure-learnability/.

Knight, Arthur (1957). The Liveliest Art: A Panoramic History of the Movies. New American Library. Available: https://archive.org/details/liveliestartpano00knig.

Kuitunen, Mika (2020). Tunks keyboards. Available (accessed on 9 March 2022): https://github.com/kulmajaba/tunks-keyboard/.

Merz, Allison, Annie Hu and Tracey Lin (2018). ClipWorks: A Tangible Interface for Collaborative Video Editing. In: Proceedings of the 17th ACM Conference on Interaction Design and Children. IDC '18. Association for Computing Machinery, pp. 497–500. DOI: 10.1145/3202185.3210758.

Morris, Patrick (1999). Nonlinear editing. Media manual. Focal Press. ISBN: 9780080512792.

Morville, Peter (2004). User Experience Design. Semantic Studios. Available (accessed on 31 January 2022): http://semanticstudios.com/user_experience_design/.

Nielsen, Jakob (1994). Heuristic Evaluation. In: Usability Inspection Methods. Ed. by Jakob Nielsen and Robert L. Mack. John Wiley & Sons, Inc. ISBN: 9780471018773.

Nielsen, Jakob (2012). Usability 101: Introduction to Usability. Nielsen Norman Group. Available (accessed on 31 January 2022): https://www.nngroup.com/articles/usability-101-introduction-to-usability/.

Nielsen, Jakob (2020). 10 Usability Heuristics for User Interface Design. Nielsen Norman Group. Available (accessed on 31 January 2022): https://www.nngroup.com/articles/ten-usability-heuristics/.

Oxford Learner's Dictionary (2022). Macro. Available (accessed on 15 February 2022): https://www.oxfordlearnersdictionaries.com/definition/english/macro_1.

Quiñones, Daniela and Cristian Rusu (2019). Applying a methodology to develop user eXperience heuristics. In: Computer Standards & Interfaces Vol. 66(103345). DOI: 10.1016/j.csi.2019.04.004.

Quiñones, Daniela, Cristian Rusu and Virginica Rusu (2018). A methodology to develop usability/user experience heuristics. In: Computer Standards & Interfaces Vol. 59, pp. 109–129. DOI: 10.1016/j.csi.2018.03.002.

Riley, Jenn (2017). Understanding Metadata. National Information Standards Organization. ISBN: 9781937522728. Available (accessed on 29 January 2022): http://www.niso.org/publications/understanding-metadata-2017.

Rubin, Michael and Ron Diamond (2000). Nonlinear: A Field Guide to Digital Film and Video Editing. 4th ed. Triad Publishing Company. ISBN: 9780937404850. Available: https://archive.org/details/nonlinearfieldgu0000rubi.

Shanklin, Roslyn, Philip Kortum and Claudia Ziegler Acemyan (2020). Adaptation of Heuristic Evaluations for the Physical Environment. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting Vol. 64(1), pp. 1135–1139. DOI: 10.1177/1071181320641272.

Sharp, Helen (2019). Interaction design: beyond human-computer interaction. 5th ed. John Wiley & Sons, Inc. ISBN: 9781119547358.

Vae, Vanida (2011). Makin Money with a mouse and a Wacom pen. Available (accessed on 18 March 2022): https://www.youtube.com/watch?v=eGnGnBUEUJo.

Zigelbaum, Jamie, Michael S. Horn, Orit Shaer and Robert J. K. Jacob (2007). The Tangible Video Editor: Collaborative Video Editing with Active Tokens. In: Proceedings of the 1st International Conference on Tangible and Embedded Interaction. TEI '07. Association for Computing Machinery, pp. 43–46. DOI: 10.1145/1226969.1226978.

APPENDIX A: CONTEXTUAL INQUIRY STRUCTURE

A.1 Introduction

Basic information: This study is related to my master's thesis on video editing control surfaces (physical interfaces or controllers).

This is a part of the initial phase where the purpose is to gather information about the work environments and habits of professional video editors

The goal of this phase is to get an idea of what kind of characteristics might be good on these control surfaces as I go towards evaluating existing products and designing a new prototype.

The results of this study will be handled anonymously, and any information related to your customers, ongoing projects etc. will be redacted from the thesis so you can speak freely about anything if you feel these things are relevant.

I will ask for permission to record audio of the interviews and take photos of the workspace for use in the research, these will not be published but I might later on ask for permission to use the photos in the publication.

A.2 Initial Interview

- Background information
  - Describe your work experience in video editing.
  - Describe your current role.
- Current work
  - What software do you use the most?
  - What other software or tools do you use on an average work week?
    - Have you any experience with editing-specific control-surfaces?
    - Have you thought about using control surfaces? (why did/didn't you choose to use one)

- Could you describe your current project, more specifically what are you working on today?

A.3 Observation

Brief the participant on "thinking out loud" and recommend using it during observation, and the possibility to discuss anything they feel is important with the interviewer.

Observe the participant working on a video project.

Pay attention to the following:

- Overall environment
  - Space
  - Lighting
  - Noise level
- Desk
  - Desk position and area
  - Ergonomics and working position
  - Objects on the table (artefacts)
- Interaction with the software
  - Main mode of input
  - Any differences in the mode of input depending on the current task

A.4 End Interview

Start with any questions left from the observation period

- Which single task takes you the most time in a day?
- What would you most want a custom control surface for?
- What do you think would make a good editing control surface?
  - Actions that should be included
  - What would you use it for?
  - What would be a natural way for you to use it? (modality and interaction type)
    * Modality: gesture, haptics, voice, gaze
    * Interaction type: e.g. gestures can be split into mid-air, mouse and touch, IMU based, foot movement, facial gestures

    * Input and output?

- What do you think makes a bad control surface?
- Looking at the following pictures of existing control surfaces, what are your initial reactions to them?
  - Pictures: Contour Shuttle Pro v2, Blackmagic Design DaVinci Resolve Speed Editor, Wacom Intuos Pro drawing tablet, Loupedeck CT

APPENDIX B: CONTEXTUAL INQUIRY AFFINITY DIAGRAM

Table B.1. *Legend for affinity diagram*

Light green	Professional editor, interviews
Dark green	Professional editor, observation
Light blue	Amateur editor, interviews
Dark blue	Amateur editor, observation
Red	Recognized themes and potential evaluation heuristics
Yellow	Recognized design considerations

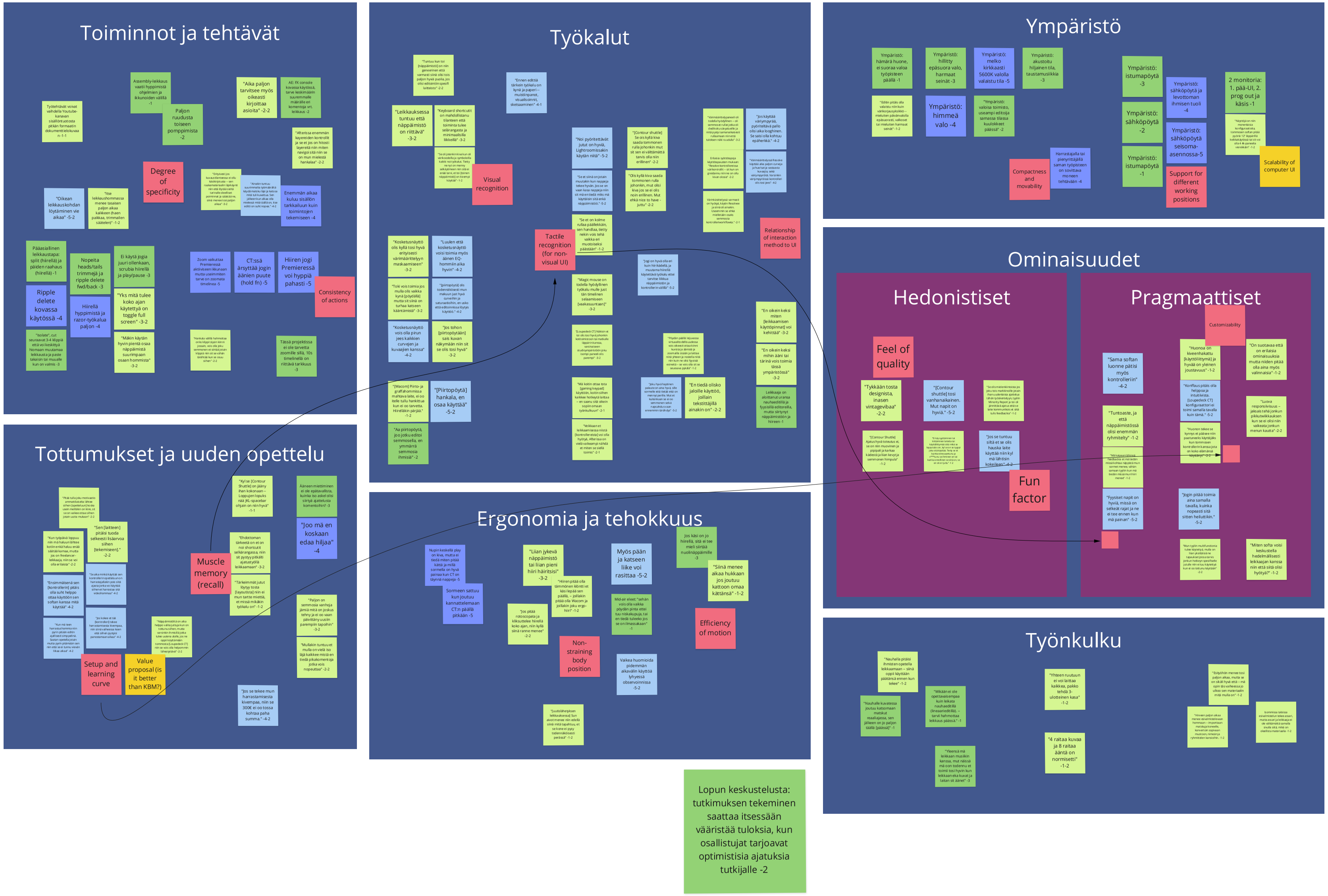
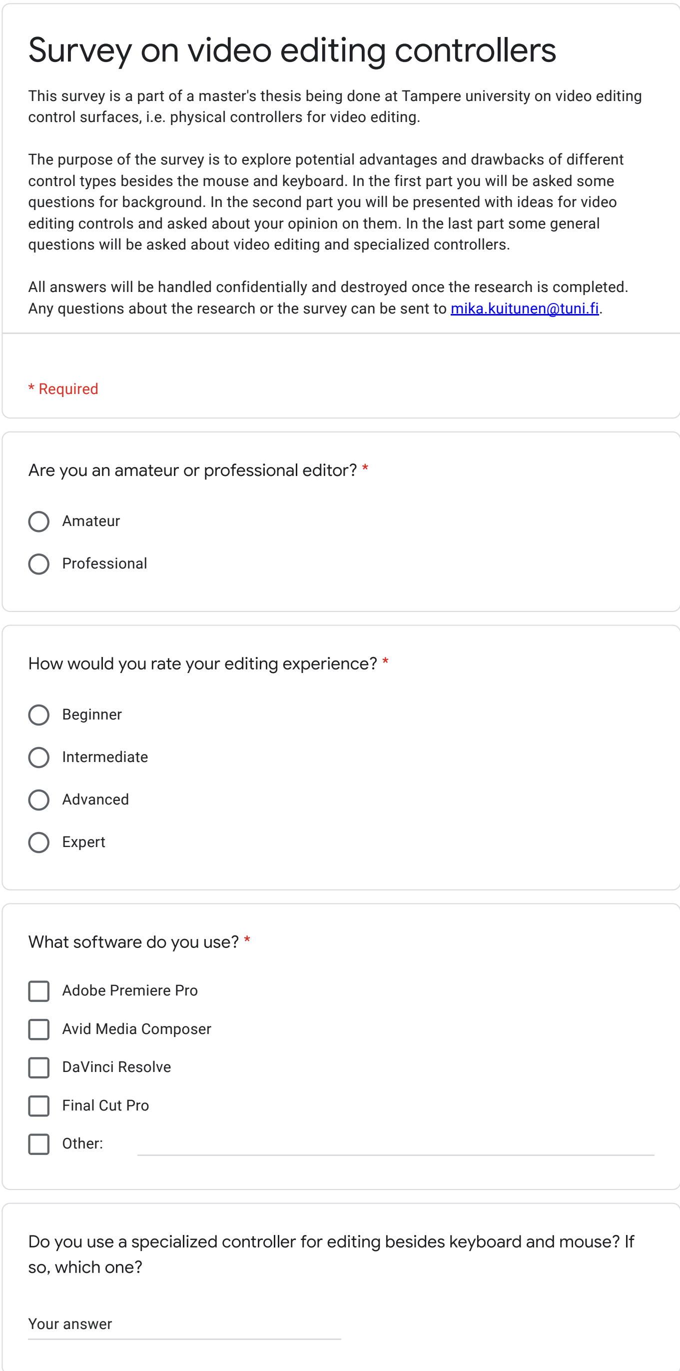


Figure B.1. *Affinity diagram for the contextual inquiry*

APPENDIX C: PROTOTYPE EVALUATION SURVEY STRUCTURE



Survey on video editing controllers

This survey is a part of a master's thesis being done at Tampere university on video editing control surfaces, i.e. physical controllers for video editing.

The purpose of the survey is to explore potential advantages and drawbacks of different control types besides the mouse and keyboard. In the first part you will be asked some questions for background. In the second part you will be presented with ideas for video editing controls and asked about your opinion on them. In the last part some general questions will be asked about video editing and specialized controllers.

All answers will be handled confidentially and destroyed once the research is completed. Any questions about the research or the survey can be sent to mika.kuitunen@tuni.fi.

* Required

Are you an amateur or professional editor? *

- Amateur
- Professional

How would you rate your editing experience? *

- Beginner
- Intermediate
- Advanced
- Expert

What software do you use? *

- Adobe Premiere Pro
- Avid Media Composer
- DaVinci Resolve
- Final Cut Pro
- Other:

Do you use a specialized controller for editing besides keyboard and mouse? If so, which one?

Your answer

Figure C.1. *Evaluation survey part 1*

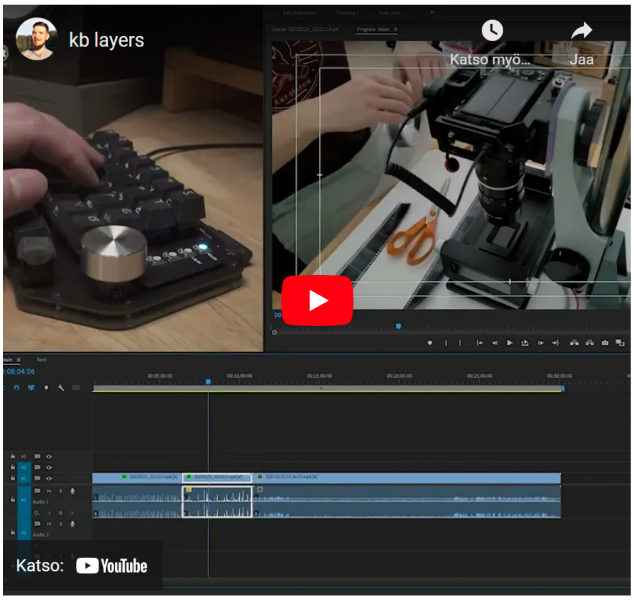
Ideas for video editing controls

In this section you will be asked about your opinions on different controls for video editing. Please read the description, watch the videos and answer the corresponding questions.

Note that even though the controller in the videos is a one-hand keypad, similar controls could be added to full-size keyboards or as standalone devices.

The controls could also be customized to do what you want, the videos are just examples on what the control could do.

Switchable layers on keyboard (same buttons for jogging and frame by frame)



Keyboard with switchable layers *

	1 - Strongly disagree	2	3	4	5 - Strongly agree
This control is interesting to me	○	○	○	○	○
This would be a useful control to me	○	○	○	○	○
I would be willing to spend the time to learn this control	○	○	○	○	○
I would be willing to pay for this control	○	○	○	○	○

Other thoughts? If you would this control, what for? What control would you use instead?

Your answer

Figure C.2. *Evaluation survey part 2, similar format continues for other interactions*

General thoughts about video editing and controllers

Please answer the following statements: *

	1 - Strongly disagree	2	3	4	5 - Strongly agree
I'm interested in specialized editing controllers					
I'm more interested in completely new control styles than reusing familiar ones					
I'm willing to spend time learning a new controller					
I'm willing to spend money on a new controller					
Keyboard and mouse are good enough for me					
I could improve my editing efficiency					
I want to improve my editing efficiency					
I'd rather spend my time editing than improving my workflow					
Editing controllers are more useful for experienced editors than beginners					

Anything else you'd like to say?

Your answer

Figure C.3. *Evaluation survey part 3*

APPENDIX D: PROTOTYPE EVALUATION VIDEOS

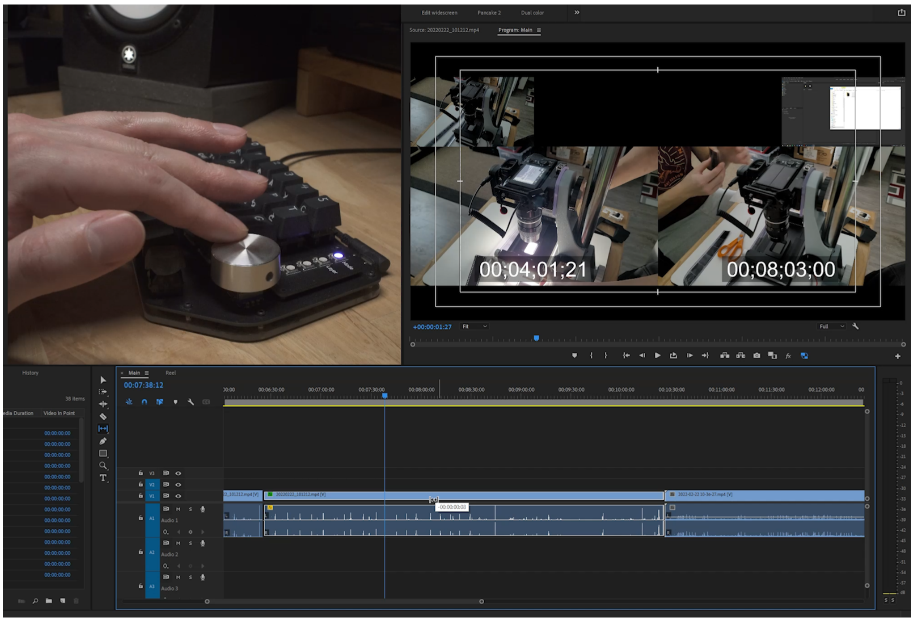


Figure D.1. *Prototype evaluation video: rotary wheel*

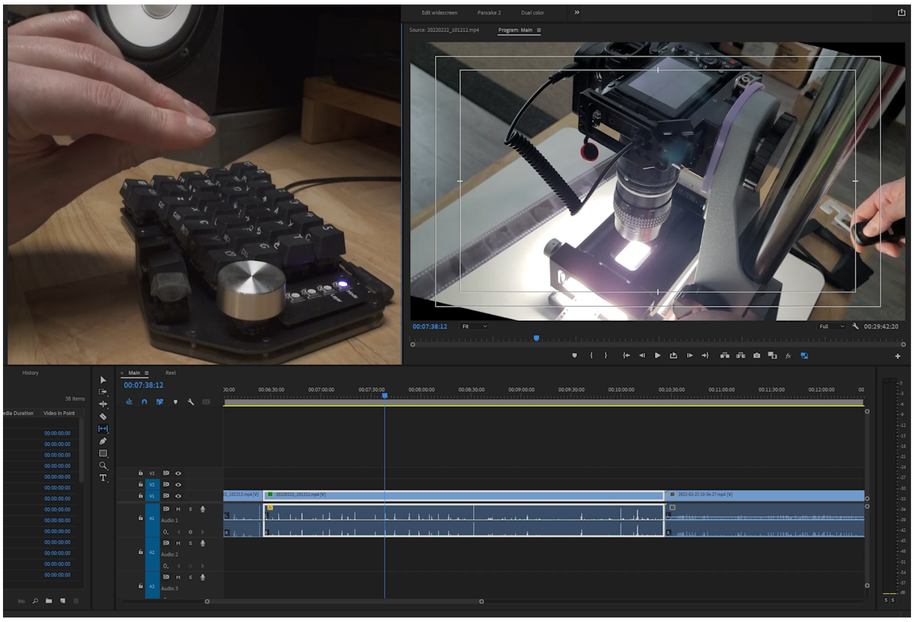


Figure D.2. *Prototype evaluation video: mid-air gestures*