
The "Black box" Problem of Teaching

("Black box" problemet i undervisning)

Simon Jøhnk Lauridsen, 201807863

Bachelor report (15 ECTS) in IT product development

Advisors: Eve Hoggan and Henrik Korsgaard

Department of Computer Science, Aarhus University

August 15, 2021



AARHUS
UNIVERSITY

DEPARTMENT OF COMPUTER SCIENCE

Table of Contents

Abstract	4
1 Introduction	4
2 Background	6
2.1 Related work	6
3 Methodology	8
4 Analytical framework	10
4.1 Presence	10
4.2 Privacy	11
4.3 Personalization	11
4.4 Fidelity	13
5 Filter Implementation	13
5.1 Sound visualization	14
5.2 Live Silhouette	15
5.3 Avatar	16
6 Analyse	17
6.1 Presence	17
6.2 Privacy	19
6.3 Personalization	21
6.4 Fidelity	23
6.5 Additional findings	25
7 Discussion	25
7.1 Discussion	25
7.2 Future work	27
8 Limitations	28
8.1 COVID-19	28
8.2 Data	28
9 Conclusion	28
10 Acknowledgements	29
References	30
Appendix	33

Abstract

With the COVID-19 pandemic remote learning tools such as Zoom have become an invaluable resource. However, they are heavily dependent on the use of cameras, which has proven to be an issue. There are several reasons people turn off their cameras during online lectures - privacy, vanity, anxiety, fatigue, etc. We refer to this as the "black box" problem. This thesis aims to explore the possibilities of using virtual webcams as filters during online lectures, in order to make online lectures an overall more enjoyable experience for the students. The focus of these filters is to improve the sense of presence, while taken issues regarding privacy into account, and the possibility for personalization. Three filters are presented in this thesis. These filters are part of a case study with seven students, who tested the filters during a mock lecture on Zoom. The experiment showed great promise in adding filter options to remote learning tools, to making sitting in front of a camera for extended periods a better experience.

1 Introduction

As of March 2020, almost every part of society was involuntarily forced into lockdown, as COVID-19 was deemed a pandemic by the World Health Organization. This quickly made social distancing a part of everyday life, which vastly changed the culture of education overnight.

Attending school in person was no longer the norm. Instead, students' study from home via various technologies to ensure that their education remains uninterrupted (Chen et al., 2021).

Live streaming presentations and video chats have replaced whiteboards and classrooms.

The students do not only rely on these platforms to connect with their teacher, but also for connecting with their peers when working together in groups.

Remote learning has become a crucial part of schooling for almost any student during lockdown.

Remote learning is not a new concept, as online courses (e.g., Khan Academy, Iversity, edX, etc) have been available for many years now in various forms. Many universities have also started to provide online courses, as they continue to grow in popularity (Giannakos et al., 2014).

However, these types of courses all benefit from having been made specifically for an online learning environment, with a curriculum made on the premise of them being taught remotely.

Most educations who were forced to adapt to online learning during COVID-19 are however not set up in such a way. The educators who were forced to adapt to this situation were not prepared to do so, and a lot of the material which is being taught is not made for online learning. This includes slides, projects, and exercises. Even learning management systems such as Blackboard and Brightspace are not made with remote learning in mind, which can lead to shortcomings. It is possible to learn from some of these remote learning tools which existed pre-COVID-19.

However, this situation presents itself with some unique challenges. In addition to being a sudden and involuntary change, students need to deal with the added stress and anxiety that comes with a pandemic. These things are some of the most noteworthy differences between remote learning from before and during COVID-19.

Research shows that there are many things to take into consideration when orchestrating an online environment for both work and education (Cao et al., 2021; Chen et al., 2021).

Nonetheless, this sudden change in medium meant that few schools had the time and resources to make a seamless adaptation into online learning. This in total makes for a vastly different experience of remote learning compared to what we have been used to up until now. This new form of forced remote learning is the basis for the research made in this thesis. In particular how the appropriation of video tools has affected students during lectures.

There are several tools available as a medium for online teaching e.g., Zoom, Microsoft Teams, Google Meet, etc. Zoom seems to be an especially popular choice for many schools, as they were able to provide a scalable service within a short time. The University of Aarhus, the organization that this thesis was made in collaboration with, has also chosen Zoom as their choice for video calls and lectures (Aarhus Universitet, 2021). Therefore, Zoom will be the basis for this thesis. Zoom is a videoconferencing software used for remote meetings, distance education, and online social relations. Zoom offers possibilities for transitioning and continuing education online despite the pandemic.

Just like traditional schooling, remote learning is not without its problems. A large percentage of students' report feeling unmotivated and stressed due to the pandemic (Browning et al., 2021), which obviously is a problem in and of itself, but this also poses a real challenge for online learning. A study among university students showed that most respondents preferred not to use webcams during online activities for reasons including privacy, anxiety, stress, and fatigue (Rajab & Soheib 2021). Some describe this problem as 'Zoom fatigue' which has become a known issue, as many people feel mentally exhausted when using video chats for extended periods of time (Bailenson, 2021). Video chat can be mentally taxing, which can lead to students choosing a more passive approach, where their presence is less noticeable. This approach might help with some issues but is a problem for remote learning. It can be hard to know which students are present and engaged during lectures when only a few (or none) participants have their webcam turned on. Student engagement is less noticeable compared to being face to face in a real classroom and getting a feel for the atmosphere is near impossible without being able to see any body language. This is what we refer to as 'the black box problem' i.e., the problem with having blank black boxes in Zoom representing students, which hides presence and engagement.

In this thesis, we explore some possible solutions to ‘the black box problem’. Changing the content of lectures is beyond the scope of this thesis. Instead, we made different changes to aspects of the medium, with the goal of making remote learning a more enjoyable experience overall. Virtual webcams were used to make filters, which are aimed to combat different parts of the problem. These filters were presented to a group of students, who got to test them out in an online lecture using Zoom.

To summarise, this thesis will discuss the following research questions;

Research Question 1, *do different representations of a media make remote learning a more enjoyable experience?*

Research Question 2, *do different representations of a media increase participants' sense of presence?*

Research Question 3, *do different representations of a media improve privacy-related issues?*

Research Question 4, *how does personalization factor into possible solutions?*

2 Background

Since the 2019 COVID pandemic hit the world a lot of people have started working and studying from home, many of which are using video chats to communicate. This increased interest in the field and the huge demand for remote solutions, led to a great deal of research into both remote work and learning. This study is built upon some of the previous work done on videoconferencing by reviewing some of that prior research and is likewise inspired by the lack of addressing “the black box”-problem explained in the introduction.

2.1 Related work

Early this year Stanford researchers posted an article called ‘Zoom fatigue’ which addressed the notion that video chatting “*exhausts the human mind and body*” (Stanford University, 2021). This article was based on a study by Bailenson (2021). His study deconstructs Zoom and evaluates its individual technical aspects.

As a result, he lists four things that he claims contributes to ‘Zoom fatigue’;

Excessive amounts of close-up eye contact, constantly looking at yourself, reduction of our usual mobility, and the cognitive load linked with video chats. He also claims that understanding these pitfalls can help build better interfaces and make users develop better habits regarding videoconferencing.

Rajab & Soheib (2021) also made a study around the challenges of using webcams during online classes. The study was based on an online questionnaire that targeted medical students at Alfaisal University, Saudi Arabia. A total of 319 medical students completed the study. They found that the main concerns among students using webcams were due to issues with privacy, the feeling of unease of having their face broadcast during classes, shyness, discomfort, and sometimes fatigue. The study also found that most students reported not using cameras when communicating via social media, and that an overwhelming majority preferred not using cameras during online classes as well. The few students who actually preferred using a webcam during learning did so for reasons such as social engagement, accountability, and mimicking classrooms. Research from the University of Erlangen-Nuremberg (Bedenlier et al. 2020) came to a similar conclusion i.e., correlation between webcam usage and feeling uncomfortable.

Chen et al. (2021) conducted a mixed-method study in which they did a large-scale analysis of live streaming (LS) based on remote education experiences in Chinese colleges. The study also explores the challenges and possibilities regarding LS learning. 30 students and 7 instructors who engaged in LS learning took part in semi-structured Interviews. Additionally, a large-scale survey study was conducted, which likewise targeted students and instructors, 6291 students and 1160 instructors in total. The study found that LS learning helps students and teachers achieve their educational goals, but with challenges such as students' difficulties in paying continuous attention, decreased learning efficacy, and lack of engagement and collaboration.

A study on web-based learning was conducted by Picciano (2002), in which he examined performance in an online course in relation to student interaction and sense of presence. The study itself and its findings are not especially relevant for this thesis, but the literary review is. His analysis of the concept of "presence" and the distinction between interaction and presence, is part of the foundation on which this thesis is built.

Sung & Mayer (2012) also did a study on social presence in online distance education. The study was based on an online social presence questionnaire given to 612 college students from two different online universities in South Korea. The study revealed five factors representing facets of social presence in online learning environments: social respect, social sharing, open mind, social identity, and intimacy. In addition to the study a definition of online social presence was explored. This likewise helped define what presence would be in the context of this thesis.

An abundance of work related to remote learning has been done amid the COVID-19 pandemic. This work done in this thesis has been parallel to many of these studies and are therefore not shaped by them. New and more relevant studies relating to this topic could have been published alongside this thesis.

3 Methodology

A case study was conducted to analyse the filters and their impact in a live setting. A case study is an in-depth examination of a few individual cases (Lazar et al. 2017). The case study was used to build understanding and provide insight into certain behaviour. This type of data collection does not result in any statistical analyses, but instead each subject is analysed with the goal of developing general principles that can help explain the motivation behind their actions. This case involved observing a mock lecture with students using the filters in a live setting. A mock lecture is not done with actual content that the professor would use in their classes. In this case the topic was “Artifact Ecologies in World of Warcraft Raiding” by Henrik Korsgaard. The lecture was done via Zoom, to mimic how lectures have been done at Aarhus University during lockdown. Seven students signed up to participate in the lecture. Each participant was instructed in how each filter worked and was presented with short videos demonstrating them all. A technical guide to set up their chosen filter was also provided¹. One participant was chosen to not use their webcam at all, two used the “sound visualization” filter, one the “Live Silhouette”, and three the “Avatar”. Eve Hoggan and I also participated in the Zoom call alongside the students and the professor (See Figure 1). The students were encouraged to ask questions during and after the lecture, like they normally would. The lecture took 20 minutes total.

After the lecture, each participant was asked to take part in a semi-structured interview (Leech, 2002) about their experience with the lecture. This interview style is mostly done with open-ended questions surrounding a framework of themes. The main focus of the questions was presence, but privacy, personalization, and to a smaller extent fidelity was brought up. The interviews included their view on the filter they were using, in addition to their thoughts on the ones they did not use. This was intended to give an insight into how the filters were perceived by both the user themselves and their surroundings.

¹ See appendix A

These interviews were recorded, which is also the case for the lecture itself. This data was evaluated by going over the recording while taking notes, and then later transcribed to use in the analysis.

To analyse the data, the recorded video interviews were re-watched, and detailed notes were taken in addition to those compiled during the lecture itself. A thematic analysis was conducted on all the notes and transcribed material (Braun & Clarke, 2012). The themes (e.g., presence) that were found were based on some of the answers given in the interviews. Based on these observations, first the themes that aligned with my research questions were examined, then the notes/quotations were grouped into those themes. Some notes fell outside any theme, these were analysed separately. This content was used to support the evaluation of the filters. These findings are presented in the analysis.

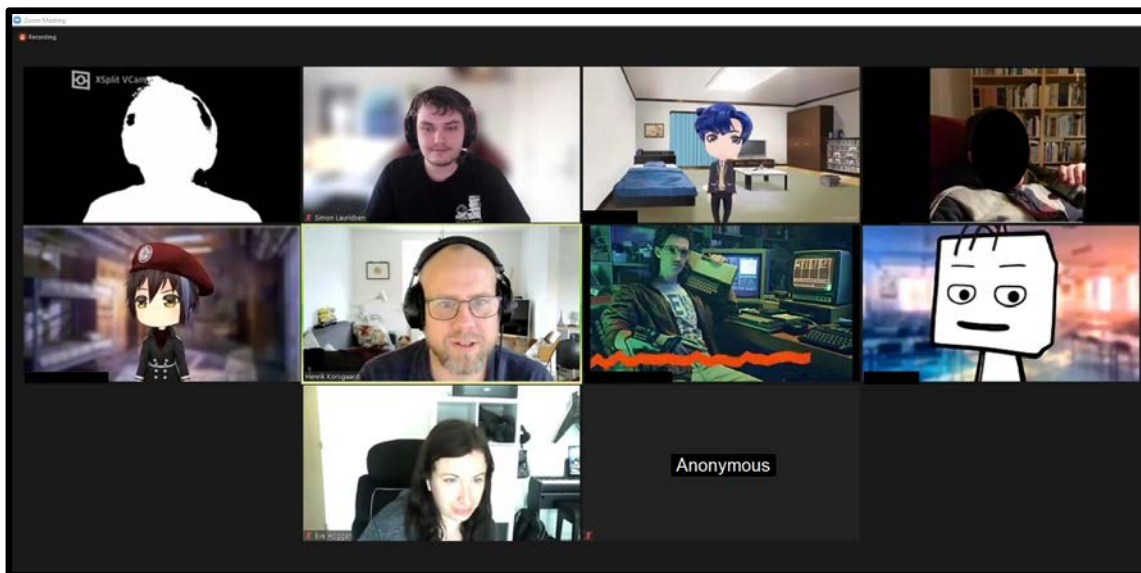


Figure 1: Every participant in the mock lecture

4 Analytical framework

In the following section, the different analytical dimensions will be introduced. These will guide both the experiment and analysis. The framework includes privacy, personalization, information, and presence (/communication thereof), with the purpose of evaluating and comparing each of the three filters.

4.1 Presence

In the context of virtual reality (VR), presence is often tied directly to fidelity i.e., Users tend to experience a greater sense of presence in a more immersive virtual environment. It is described as *“the psychological sense of “being there” in the environment generated by the system.”* (Lee, 2010). This definition is too narrow for this project, as low fidelity does not necessarily dictate a low sense of presence. In the context of this project, presence (or the sense of it) can still be experienced by someone using a low-fidelity communication option. Furthermore, presence is not only tied to a single user and their experience, but also to how others perceive their presence.

The topic of teacher-presence (/lecturers-presence) is also something related to online learning. Teacher-presence is often associated with connectedness, which is achieved when the teacher is putting in the effort to be involved and engaged in the learning (Stone & Springer, 2019). In short, teacher-presence is mostly tied to how well-supported students are feeling during online learning, even though being able to literally see the teacher is relevant to this type of presence (Hehir et al., 2021). This project tries to achieve a similar connection, but instead with student-teacher presence and student-student presence. In other words, this thesis focuses on students having a sense of being present themselves,

Social presence is another notion often used to describe presence in online learning. Social presence has had many definitions over the years among others;

“A student's sense of being in and belonging in a course and the ability to interact with other students and an instructor although physical contact is not available” (Picciano, 2002),

“an individual's ability to demonstrate his/her state of being in a virtual environment and so signal his/her availability for interpersonal transactions” (Kehrwald, 2008),

and recently *“the ability to perceive others in an online environment”* (Richardson 2017).

All three definitions align with how presence is used in this thesis. A sense of presence is experienced when students can grasp each other and have a feeling of ‘being there’.

In summary; Presence in the context of this thesis will resemble what might usually be defined as social presence. Presence is not only how immersed someone is in the system, but also how well others are able to sense your attendance.

4.2 Privacy

Most people who were forced to work/study from home were not prepared to do so. Some people do have offices, but that is not the norm especially not for students. This means that a lot of people now must share some of their private space with others when using a webcam at home, something that might not be a problem during online social events but can be problematic in a more formal setting. Privacy in the context of this project is therefore tied to how much of one's personal information is shared with others. This does not only include one's home but also one's person, as appearances might be different for someone at home than what you would expect from one going out in public. It is unreasonable to expect the user to have control over all personal information within the view of the camera (Anwar, M. 2020), therefore one might unknowingly share personal information when using a webcam.

Potentially exposing one's home is a concern reported by many students using webcams for remote learning, and privacy concerns are the main reason for not wanting to use webcams in online classes (in addition to anxiety) (Rajab & Soheib 2021). In addition to all this, cameras are often associated with the loss of privacy in general (Rajpoot et al., 2015), which makes the use of webcams in online learning a privacy concern in and of itself.

In summary; privacy is about not sharing unnecessary parts of the user's personal life with others, being either their appearances or the (personal) space they are in. This is what will be considered in this thesis.

4.3 Personalization

Appearances are important, people want to personalize everything from clothes to their phone background. Online environments are no different. This is especially common with communication apps, such as Facebook, Instagram, Snapchat, etc. The way people customize their experiences can be separated into two categories: Private customization and social customization (Griggio et al., 2019).

Private customization options are only visible to the user who made them, in other words, it is the settings the users change to make the app function as they want it to.

Social customization options are visible to the user and their surroundings and allow them to express aspects of their identity. In addition, social customization also lets users express themselves differently to different social circles within the same system, as one would in real life when talking to different groups of people e.g., colleagues, friends, family.

Zoom does allow for a wide variety of options for private customization, as one would expect for a modern piece of software. The options for social customization are however limited to, changing profile pictures, using virtual background, and adding visual gimmicks to one's webcam feed.

Another way of describing personalization is that of a dissatisfier in the two-factor model of work satisfaction (Hassenzahl, 2010). Having no options for personalization reduces satisfaction, but appropriate or even disproportionate many options do not likewise guarantee high levels of satisfaction. This is also described as the hygiene factor i.e., bad hygiene is bad, good hygiene is neutral/expected. Personalization makes the difference between a dissatisfied and neutral user experience. A communication app without proper options for customization and personalization will be perceived as dissatisfactory to the user, making it an essential aspect of the system.

Another critical aspect of the psychology of personalization (/customization) is control.

Users want to have control over the performance of the system they are using (Hassenzahl, 2010; Marathe et al., 2011). Having too few options for controlling a system will likewise leave the user dissatisfied.

Personalization is also regularly tied to customization. How the user interacts with an application (e.g., key bindings, shortcuts, menus, etc.), and visual preferences (e.g., fonts, colours, sizes, etc.) are often both part of customization (Mackay, 1991). It is common to differentiate between appearance-based or functionality-based changes, but still considers both customization activities (Marathe et al., 2011). However, in this thesis, those two are kept separate and the focus will only be on the appearance-based changes that others can see i.e., the things that users can change to express their identity to others e.g., profile picture, backgrounds, colours, etc.

In summary; In the context of this thesis, personalization is defined as appearance-based customization. Meaning it is the changes used to express identity and set yourself apart from others within a program.

4.4 Fidelity

Fidelity is a common topic within the context of VR, where it is often used to describe how accurately an effect is reproduced within a system, or “*the visual consistency between the real-world and virtual objects*” (Thompson, 2020). This means that a high-fidelity VR system is one that creates a virtual experience like that of a real-world experience (Han et al., 2020). This way of describing is not in line with its use in this thesis, as mimicking the real world is not a focus of the filters.

In communication theory, fidelity is often referred to as bandwidth (Nardi, 2005). Bandwidth is how well a media allows for the exchanges of information. High bandwidth channels for example video chats, allows for a substantial amount of information to be shared, such as body language and other types of nonverbal communication. On the other hand, low bandwidth channels such as text messaging only allow for little information to be exchanged. This way of evaluating the exchange of information is what fidelity will describe in this thesis.

In summary; Fidelity in this project is defined by how well a person is represented to their surroundings through the medium. A lot of social cues are shared when people meet face to face, such as body language and types of nonverbal communication. Higher fidelity communications provide most (if not all) of these things, where lower fidelity communications only produce a few aspects. This distinction between fidelities is what is being analysed for each of the filters.

5 Filter Implementation

Three different techniques/solutions were identified and used to conduct experiments with presence in Zoom meetings with. In the following section, I will describe each of the filters, their purpose, and how they were implemented. A more detailed explanation of how to set up the filters can be found in Appendix A.

All three filters create a virtual webcam, meaning they take some input, and output a fabricated webcam source to be used in programs (See Figure 1). Most of the filters were created using OBS Studio², which is a popular software choice for video recording and live streaming. However, OBS can also output any recordings as a virtual webcam, which was useful for implementing these filters and doing the experiment. Additionally, another program, Xsplit Vcam³, was also used in the creation of the filters. Xsplit is a webcam background removal software, which allows for background blurring, removal, or replacement without a green screen.

² Open Broadcaster Software. Retrieved August 13, 2021, from <https://obsproject.com/>

³ Xsplit Vcam. Retrieved August 13, 2021, from <https://www.xsplit.com/vcam>



Figure 2: Virtual webcam setup

5.1 Sound visualization

The first of the filters was a virtual webcam with a sound visualizer layered on top of a profile picture of the user's choice (See Figure 2). This was achieved by using OBS along with the Spectralizer plugin⁴, which allows the user to create an audio visualization of OBS audio sources. First, an arbitrary background was put into OBS. Then the Spectralizer was layered on top, using the input from the primary microphone. OBS is then able to output this as a virtual webcam. This creates a picture (like that of a profile picture) with a moving line on top, which is a visualization of the sounds picked up by the user's microphone. This is independent of whether the user is muted/unmuted in Zoom, meaning any background noise, breathing, or loud keystrokes will be visualized to some extent depending on the sensitivity of the user's microphone.

This filter is the closest to the black box. The presence of the user is indicated by the aforementioned moving line. This is meant to inform others that something is happening on the other end, even if the user is not present in front of the PC. This lets the user keep almost the same level of privacy as they would otherwise have by sticking to the black box, but with the added benefit of being able to communicate their presence to their surroundings. However, what the user is doing and whether they are paying any attention or not to the call cannot be seen. Furthermore, not all moments of the line are caused by the noise made by the actual user, which leads to a lot of false positives. These false positives are not a problem, as the filter is only meant to show others that there are some activities, not what it is specifically.

⁴ Spectralizer. OBS Forums. Retrieved August 13, 2021, from <https://obsproject.com/forum/resources/spectralizer.861/>

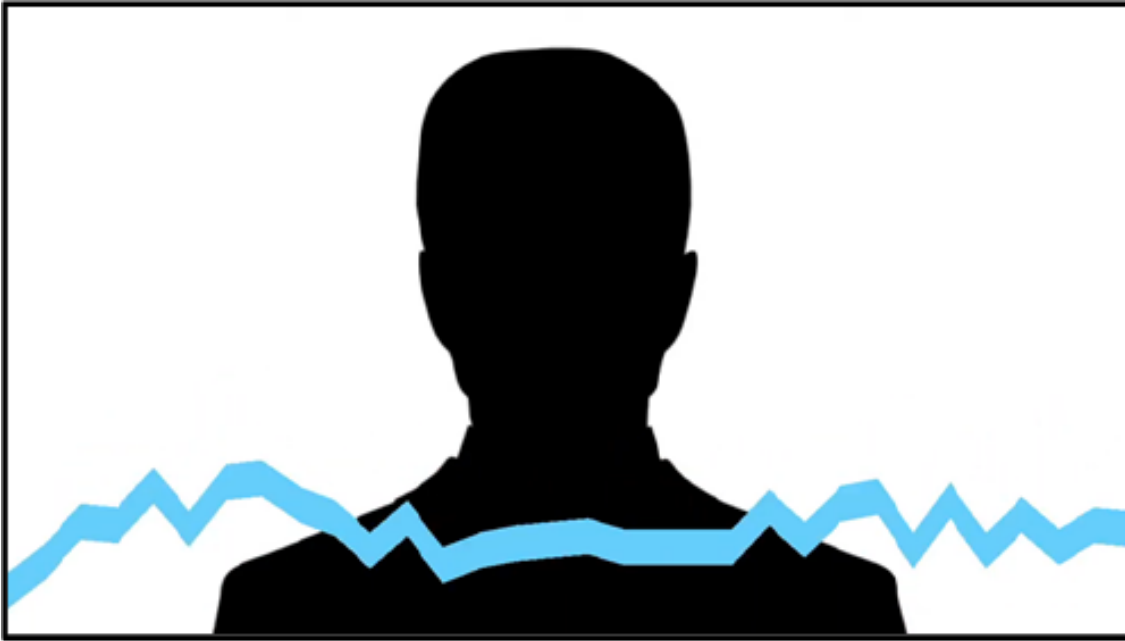


Figure 3: The sound visualization filter

5.2 Live Silhouette

The second was a live silhouette filter which turns the outline of the user into a figure with no distinctive features, which can be layered on top of a background of their choosing (See Figure 3). OBS was again used for this, along with Vcam. Vcam was used to make a form of digital green screen, by outputting a virtual webcam with no background. This was then inputted into OBS, on top of an arbitrary background. The character (which was outputted by Vcam) was then changed into a solid colour (in this case black). This outputs another virtual webcam (again from OBS), which is like a living silhouette of a person.

This filter is meant to inform the other participants of the literal presence of the user but without any associated features. This is meant to show the fact that the user is in fact there while attending the call, but without the user needing to worry about their appearance. The intention is to increase privacy, while allowing the user to still be able to use their body to make signals, such as raising their hand, giving thumbs up to a message, or showing that they are not back from a break by physically not being there. The filter does not allow the user to hide completely from the call like with the previous one, but they are still not disclosing any disinterest due to their facial expressions being covered up.

The intent is that this filter will display the user's presence to a greater extent compared to the sound filter while remaining more private than with a live webcam with no filter.



Figure 4: The silhouette filter

5.3 Avatar

The third was an avatar filter. This turns the user into a virtual character which in real-time mimics some of the user's characteristics (See Figure 4). There are multiple software options for this purpose. Some of the higher quality needs to be purchased, I instead went with a free option, as I wanted to encourage potential participants to participate in the project. I chose VTuber Maker for the experiments which are available for free through Steam, the video game digital distribution service. The software lets the user choose from a handful of pre-rendered avatars along with multiple backgrounds. Once the avatar and background are chosen, the program will mimic some features of the user, such as head tilt and some facial features, along with audio-based lip sync. This is then outputted as a virtual webcam like how OBS does it and can be used in the same way.

The filter is quite different from the other two. The avatar filter does not reduce the features of the user to just sound or a simpler image. Instead, it changes it completely. The presence of the user is therefore tied to the information provided by the avatar. This includes head movement, blinking, and mouth movement, for the software chosen for this experiment. Arm and upper body movement are not tracked, meaning a lot of body language is lacking when displaying the user's presence. Most features of the user are hidden, leaving only some of their facial expressions on display. This makes the user almost completely anonymous when using the filter, while still being able to present a character through whom they are displaying their presence.



Figure 5: The avatar filter

6 Analyse

The following section will present the findings gathered from the experiment described in Methodology. A picture of the entire Zoom setting can be seen below (Figure 5).

For the sake of anonymization, the names of the participants have been changed to their respectable filter e.g., Silhouette 1 or Avatar 2. All the interviews were done in Danish; therefore, all citations have been translated to English for the purposes of readability.

6.1 Presence

Presence was a major theme in each of the interviews, as it was the primary goal of the filter.

Many interview questions were for that reason focused on presence. The other three topics also had some overlap with presence which sometimes led the conversation back towards this topic.

It is clear from the interviews with the students that they wanted to show their presence while attending online classes in one way or another. The students wanted to show that they were attending and listening, even though they did not necessarily were using their webcam.

Some of the students also noted that it could be demotivating to present in front of black boxes on Zoom and that it would be nice to be able to communicate their presence in a better way.

As for regular Zoom use, presence is mostly experienced through people turning on their webcam or using the chat/unmuting to participate directly.

“There is a quality in reflecting that you can still be present, without physically being in front of the monitor. [...] I do not leave the lecture just because I go and brew a pot of coffee, but if you watch the webcam then it can feel very absent if you just disappear for two minutes, even though you are still listening via headphones. I thought it was nice with the sound feedback, at least if you have a wireless headset, so that you can show your presence whether you are sitting in front of the camera or not.” (*Sound visualization 1*)

Most of the participants thought the sound visualization filter to be useful and a good way to communicate their presence. Even though this filter made the smallest change, it was still perceived to be a decent add-on. One participant thought that it should be the default in Zoom instead of the current black box, as it added some feeling of presence without requiring much effort by the user.

However there were some concerns, some of the participants mentioned that since you are silent most of the time during a lecture no sound will show on the filter, and any sound shown is more often than not due to background noise or similar. In other words, the presence displayed by the filter did not feel genuine to some users and could be perceived as fake/fabricated.

“[...] I could easily imagine that I would use that as a solution, as it is fully anonymous but with me still being present with movement.” (*Silhouette 1*)

The live silhouette filter was the most popular filter in terms of presence. This was not unexpected as it is the closest to just using a webcam as intended. Most of the users felt like they were present at the lecture, but one user was a bit reluctant. The user noted that since the person(s) using the filter did not move much, it just felt like a static picture of a silhouette, which did not do anything for them in terms of presence. Another user also noted that the filter might feel a bit ominous if it were to be used by an entire online class at once. Both things could be addressed in future work.

“I actually liked it. I felt the same way about it, as I felt about the silhouette, it is just a different way of doing the same thing.” (*Visuel Lyd 2*)

Feedback on the avatar filter was more unresolved. Some of the participants felt that the people using the avatars were in fact present during the lecture and that it was usable for anonymous participation. However, others felt different and thought the filter had some glaring issues. Despite being rather advanced the avatar filter is still rather low on fidelity. In other words, the number of details captured by the avatar software was too low for some users, which took away from their experience of presence. Another user noted that the avatar is also not disabled itself when the user moves out of frame, this can also give a false impression of presence when using the filter.

In summary; Students want to communicate their presence to others around them when attending classes online. The regular Zoom does not offer many options for users to signal their presence. This makes for a binary experience where the user is limited to either having their webcam turned on/off, which in turn makes every new addition a welcome one.

All the three filters seemed to provide some sense of presence for the students but to a varying degree. Despite only making a minor change to the user's profile, the sound visualization filter did make noticeable changes to the communication of presence when working correctly. The silhouette was most well-liked and seems to have the most potential, as it was the closest to regular webcam use, but without the drawbacks. The avatar was well received by the user but would be better with higher fidelity e.g., picking up on more kinds of movement.

6.2 Privacy

Privacy was generally a concern for the students primarily perceived privacy. Extended use of video chat can evoke the sense of being watched and loss of privacy by those who are in the view of the camera. Not having to deal with this seemed to be a universally appreciated aspect of the filters. No comments related to privacy were made on the sound visualization filter, but this is most likely because this filter is so close to the status quo, and privacy is close to maximum when using the filter. However, since the filter does share all audio inputs some privacy is theoretically lost as others could monitor this. Even though the filter only functions as a sensor and does not distinguish between background noise and spoken words, some users could experience it as a break in privacy regardless of the actual nature of the filter.

“You do not have to think about people sitting and looking at you while you pick your nose or whatever. [...] People can see the outline of you being there. You can show that you are there and that you are present, without feeling that you are on display.” (*Visual Lyd 2*)

The silhouette however does improve privacy significantly, as one’s physical appearance is covered up. This was appreciated by most of the students, as it took away some of the concerns related to looking presentable when using one’s webcam. This also makes it so that the user spends less time looking at themselves, which can be an annoyance for some. However, this was only commented on by a single student, who found it to be a nice addition.

“I cannot stand up in a serious classroom where we do linear algebra, computer architecture, or similar, and share my webcam with this being my room [anime poster in background]. Even though I might find it to be something funny and nice, it is still something that does not belong in a professional context. The avatar solution might be a smart idea for solving these things, both with background and personal aspects, as it is less your own self that can be judged by others, but instead your avatar.” (*Avatar 1*)

This quote exemplifies a user who wants to show his attendance and presence but find the binary nature of Zoom displeasing. He also introduces an interesting nuance; Privacy is not limited to the individual and one’s person, but also includes one’s private space. Sharing one’s hobbies, personal belongings, and similar can be a private affair that the user of video chat might unwillingly participate in. This is also a distinction between what is professional and non-professional. Offices and classrooms are meant to be professional, while one’s personal space such as a bedroom is not. Having a home office is a rarity among students, and many are constrained to living in just one room. Having to share this room via video can create embarrassing moments for the user, as a personal/private space is often not meant to be professional. This is also pointed out by the user above, as this seems to be what he is experiencing.

The avatar filter shares some of the same qualities as the silhouette, as the user no longer has to worry about their looks, which seemed to be appreciated by some. On top of that, the filter removes the background and thereby the need to have a presentable room, this was especially important for two of the students. One of them was concerned about privacy in general and liked the option to blur or remove his background when using his webcam. The other, as mentioned earlier, felt like his room was unprofessional and unfit to showcase, and was therefore happy to change it into something else entirely.

In summary; Privacy was a concern for most of the students for various reasons. This included privacy at the individual and bodily level, but also privacy in terms of private space. Some of the users were concerned about looking presentable and professional, while others were troubled by the setting they had to present in front of the camera.

The current Zoom setup has the user choose between total anonymity and being fully on display. This might discourage some students from participating. Instead, the three filters add options in between for a more balanced approach to privacy, which seemed to be appreciated by the students. Each of the three filters has varying degrees of privacy (and perceived privacy) and can therefore be used in different scenarios depending on the user.

6.3 Personalization

As explained in the analytical framework, personalization is important for the user to feel as individuals, and in control of the system they are using. Without being able to personalize each different filter, students might be hard to recognize from one another. Zoom allows for personal profile pictures in addition to name tags when not using a webcam, which helps with recognizing people in meetings/lectures. When using a webcam, the user instead has the opportunity to use a virtual background which can either be a personal picture or an animated GIF.

The sound visualization filter does not provide much in terms of personalization, but both participants did use it with a profile picture of their own choosing instead of the default picture. It is hard to attribute any value to this but not everyone has a profile picture on Zoom, so this filter could encourage more people to customize their profile.

When designing the silhouette filter, the anticipation was that it would be rather restrictive or limited in terms of personalization, but this was proven wrong by the student who used to filter during the experiment. The setup used during the lecture was simplified to avoid distractions, but other variants were showcased before the lecture started and again just before everyone was split up for interviews (See Figure 5). This proves that the filter allows for more personalization than first assumed and that OBS has more uses for virtual webcams than what can be explored in this thesis.

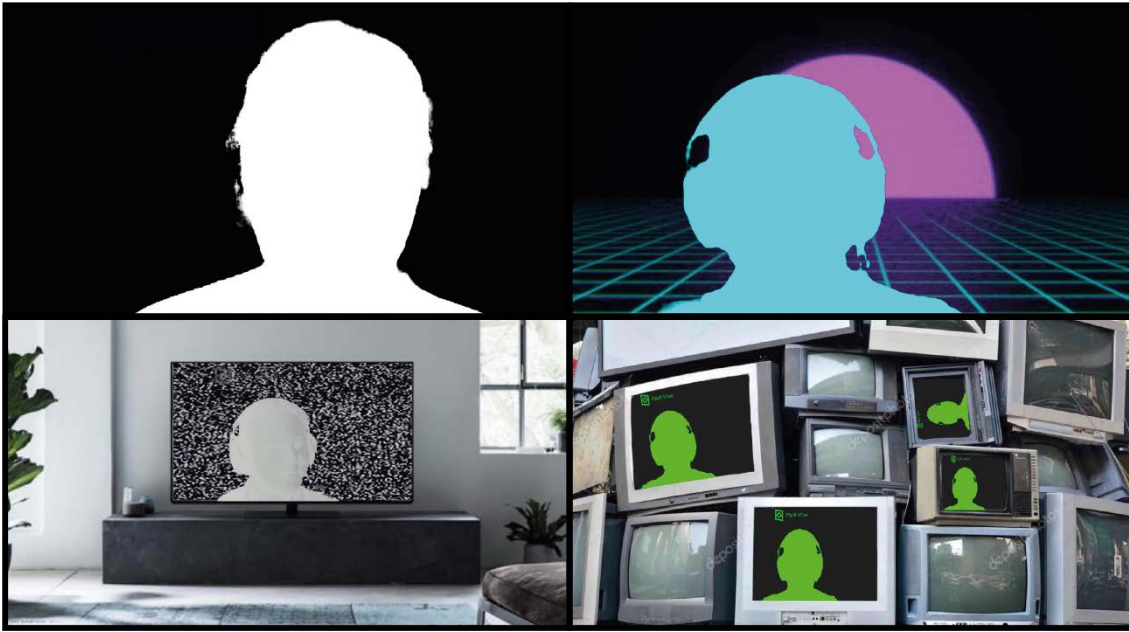


Figure 6: Silhouette variants

“[On selecting an avatar] I thought about going all out, but I think that would have been too silly. Instead, I tried to look for something more harmless. It is probably better to choose something less weird, especially since Henrik will be my examiner later
Laughter. [...] I think the avatars are a great idea, as they make room for personalization. That is, we can be individuals, we can be a bit different.” (*Avatar 1*)

The avatars were the most creative, as lots of different options for personalization are available when selecting an avatar in the software (See Figure 6). All the three users chose different avatars and personalized their experience to look unique.

The users mentioned that avatars would not function as well if multiple people were to use the same character, as it could create a lot of confusion if two people look the same. This could limit the size of the gathering that can use the filter, but this is highly dependent on the software that is used. This issue would be less likely to occur if the user had to create their own avatar upon use. It was also mentioned that if the avatars were to be used as part of an online classroom, a more focused art direction would be appreciated. The software used for this experiment had a vast variety of avatars available with little to nothing in common. Having a coherent art style for the avatars would also solve the most glaring issue noted by almost all of the students; Some of the avatars available in the software used are inappropriate and unprofessional. The students who used the avatars all noted that they specifically avoided those and instead looked for a more inoffensive choice.



Figure 7: Avatar variants

In summary; Personalization is an essential part of the Zoom interface and needs to be taken into account when adding any additional tools such as the three filters.

OBS which is used to create two of the virtual filters allows for close to unlimited possibilities for customization of the filters, which was displayed by the silhouette filter user, who displayed this to a great extent by creating multiple creative variants of his filter.

The avatar allows for much personalization but this is dependent on the software used. It is crucial that there are enough options available for customization to make every avatar feel unique and personal. This was not a problem with just three students using the filter, but it could become an issue if an entire class at once were to use the filter.

6.4 Fidelity

Each of the three filters has varying degrees of fidelity, meaning some share more information about the user than others. The sound visualization filter has the lowest fidelity, except for the black box, as it only shares sound via visualization. The avatar is at the medium level but can be changed to a higher level of fidelity with better software. The silhouette filter has the highest fidelity, as it emulates nearly all elements of regular webcam use, which has the highest level of fidelity within Zoom.

“There is not much sound coming from someone who is watching a lecture, whether it is over Zoom or not, so I think a lot of what is being visualized will be random noise, which does not really indicate that there is a person there.” (*Avatar 2*)

As mentioned earlier, the fidelity of the sound visualization filter was a bit hit or miss. Some users felt like the feel of presence was lacking, as the visualization was often either based on random sounds or not there at all, while others really liked the visual representation of giggles and similar. To increase fidelity a different/an additional sensor could be used instead of just the microphone.

“But I think, if I were to do this in the future, I would go with the silhouette. The avatar was there all the time, it is still there if you go out of frame. I felt like the silhouette was better at showing that a person was sitting there.” (*Avatar 2*)

The silhouette filter had the highest fidelity of the three filters. The primary complaint of the other two filters was that they lacked fidelity which hurt the perception of presence. This could be why this filter was liked by most of the participants. The filter captures almost all the same information as regular webcam use, but with fewer details, which is why the fidelity is so high compared to the avatar which only captures a limited amount of information. A couple of students did note that when sitting still the silhouette would sometimes look like the default picture in Zoom. This can easily be solved by customizing the filter ever so slightly, some examples of customization can be seen above under personalization.

“They were not quite “animated” enough, it might as well have been programmed to do randomized movements. I did not get the feeling that there was a person sitting behind it.” (*Visuel Lyd 1*)

As previously mentioned, the avatar lacked fidelity. Even though it seemed like an avatar filter would mimic a person rather well, the software used did not pick up on enough details for it to be the case. All the students agreed that the avatar would function better if it were to pick up on more movement. This can most likely be solved by using better virtual webcam software. As explained earlier this software was chosen due to it being free and easy to use for participants, but future work could investigate other options.

In summary; The three filters range from a lower level of fidelity to a higher one. This makes for a fine addition to the current Zoom interface which only provides the zero-fidelity black box and the maximum fidelity use of a regular webcam. However, feedback from the users does show that all three of the filters could be tweaked to better represent the information they are meant to share.

6.5 Additional findings

In addition to the findings outlined in the four themes above, other notes were taken which did not fit into any particular category. The first thing observed was that every single participant was excited by the filters, especially their own. Despite being a clear signal that choices are anticipated in general, this can also be interpreted as an unarticulated general dissatisfaction with the current options within Zoom. Such enthusiasm probably would not be displayed if the status quo was satisfactory.

Secondly, most participants seemed to like the silhouette, and a couple even said they preferred it over the one they were using themselves. This speaks well for the silhouette filter and its potential, which seemingly makes it the most promising of the three filters.

7 Discussion

In this section, a summary of findings of the study and the associated implications for the design are presented and discussed.

7.1 Discussion

This study illustrates how virtual webcams can be used as a tool for improving online learning experiences and create a sense of presence while tackling some of the issues with ordinary webcam usage. In the following section, the four research questions posed in the Introduction will be discussed.

First, do different representations of a media make remote learning a more enjoyable experience?

Using virtual webcams in an online learning setting was a modest success, as feedback from the participants in the case study was overwhelmingly positive. This does not necessarily speak for the quality of the filters, as feedback suggested that they all need refining in some way or another, instead it points to a lack of good customization options within the current Zoom interface. Adding more filter options for users could make for better experiences with online learning in the future.

Second, do different representations of a media increase participants' sense of presence?

The answer seems to be yes. The current Zoom interface is rather binary, and only allows the user to show their presence by turning on their webcam or having it disabled which greatly reduces the sense of presence. The three filters proposed in this study provided a range of new options for the students. Each filter provided a sense of presence but each at a varying degree. In other words, the filters did provide some sense of presence, some more than others as intended. Having more than just the simple choice of turning the camera on/off, was appreciated by the students, and should be considered as an inclusion in interfaces like Zoom in the future.

Third, do different representations of a media improve privacy-related issues?

Privacy was a common concern among the students when using their regular webcam during lectures. This included both sharing their personal space with others, and concerns about looking presentable in front of other students. Zoom does have the option for users to blur/change their background when using a webcam, but this only solves half of the issue. This made the ability to remain private while communicating a sense of presence a praised aspect of all three filters. In other words, having the option to choose between varying degrees of privacy should likewise be considered in the future if more filters were to be added to Zoom.

Finally, how does personalization factor into possible solutions?

As explained in Methodology, personalization is an important part of online environments Zoom included. The user wants to be able to express of their identity and distinct themselves from others. Zoom does have some personalization options, but the filters allowed for so much more creativity, which was clearly seen during the lecture. Personalization was not a topic that many of the participants commented on, but everyone did change their filter to some unique, despite being given the same instructions. This means that some of the students chose to explore OBS on their own, which reaffirms that there is the wish to express uniqueness. In other words, this study confirms the importance of personalization and the option to express identity, which should be considered when developing online learning environments.

7.2 Future work

Future work can be taken in many different directions. The most apparent would be to do a bigger experiment with several lectures and more students. Interviews with the lecturers/teacher are also an additional resource for evaluation which could be added in the future, as their perspective could give insight into how the filters are perceived by the one who is carrying out the meeting, lecture, etc.

OBS was an invaluable tool in the creation of the filters, and it became clear during the development that the ability to create virtual webcams at will had immense potential. OBS is a well-explored piece of software for content creation with many other uses. OBS is an especially popular tool on streaming-platforms such as Twitch and YouTube, where it is commonly used for screen capture. Using OBS in an educational setting has yet to be explored. People are already using OBS in various ways, so involving students in the development of filters and/or as input to creative and useful uses for filters, could be explored in future work. The students who participated in this study displayed great creativity, so having the students design their own filters could lead to otherwise unimagined solutions and affordances. None of the work related to this study seems to include students in the design of online learning tools like Zoom but including students could prove to be a valuable resource, especially since OBS is familiar to some students already.

The implications of the design seem to suggest that adding more filters to interfaces like Zoom is appreciated by the users. However, implementing new filters into Zoom could take a substantial amount of time and effort, as the filters would most likely have to be redesigned and improved upon to fit the current Zoom interface. Another option could be instead to introduce students to how to create their own filters. Since OBS is already available for free today, all it would take is a short introduction and/or guide into how to use virtual webcams. Any institutions who want to use custom filters can implement them with short notice, if the COVID-19 pandemic (or something similar) were to force people into lockdown for an extended period of time yet again. This might be relevant in the future and should be explored by any university looking to improve their online learning experience.

In summary; Creating a new medium for videoconferencing is time consuming and frankly impractical. Instead, this study suggests future work revolve around adding tools such as webcam filters to already existing platform.

Future designs should also include students in this process, as they have shown great creativity.

8 Limitations

This thesis had some limitations and restrictions, which will be discussed in the following section.

8.1 COVID-19

Despite being the foundation of this thesis, COVID-19 has also been a hindrance to the project. It is hard to say how big the impact has been but working from home during lockdown has most likely been stressful in some way or another for everyone involved with this project. Every meeting had to be done remotely, which meant less flexibility in supervision overall. This also meant that recruiting students for the experiment had to be done via email, which resulted in no replies the first couple of tries, and only a few the third time. The low number of replies might also be due to added stress from COVID-19 among students, who were already busy with their own projects. Recruiting participants by presenting the project and talking face to face, would probably have been more effective, and could have led to more participants.

8.2 Data

As mentioned above, the number of participants who signed up for the experiment was rather low. This resulted in a limited amount of data available for analysis. Nonetheless, this meant that it was possible to do interviews with every single participant, which provided valuable evaluation. No of the data was lost, but an error was made when recording; My own voice was recorded neither during the interviews nor the lecture itself, meaning the recorded interviews only had the answers from the participants and not the questions asked. This made the analysis a bit harder, but I was able to piece together what questions were being asked in each interview based on my memory and written script, despite the interviews only being semi-structured. I do not consider this to have had a meaningful impact on the final result.

9 Conclusion

Despite its shortcomings Zoom has undoubtedly been an indispensable communication tool during Covid-19. Students, teachers, and others alike have all benefited immensely from having good reliable software, which has made the transition to remote learning/work possible in the first place. This thesis takes on some of the known issues with the current Zoom interface design, by adding custom made virtual webcam filters on top.

Seven students participated in a case study, in which they engaged in a mock lecture using the filters outlined in this thesis. Semi-structured interviews were conducted with the students afterwards, and the findings was evaluated in a thematic analysis.

Overall a filter-based approach to increase presence and decrease 'Zoom fatigue' in online learning environment seems promising. The students were excited to use the filters and could see themselves using something similar in the future. In short, using the filters to improve the experience of online learning was a modest success, and issues regarding both presence and privacy could most likely be solved using this technology. However, it is hard to derive any clear-cut evidence from having only tested it in a mock setting with so few students. Hence, future research that details virtual webcam use in an educational setting and includes student perspectives is needed. Research opportunities are plenty, and hopefully future studies will test these ideas.

10 Acknowledgements

Thanks to Eve Hoggan and Henrik Korsgaard for their valuable supervision, including suggestions and feedback. An especially huge thanks to Henrik who dedicated a lot of this time to regular meetings and was a big help in recruiting participants. I am also grateful to the students who participated in my experiment, without whom this thesis would not have been possible.

I also want to thank some of my friends -

Simone Randrup for being so supportive and an all-around great friend, Jonas Mertz for keeping me entertained during the lockdown, and Erik Poulsen for assisting in various ways.

Finally, I wish to thank Aarhus University for letting me do this project and providing a student welfare counsellor to help me cope with the stress and anxiety related to the COVID-19 lockdown, without that guidance this project would not have been possible.

References

- Ai-Lim Lee, E., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education*, 55(4), 1424–1442.
- Anwar, M. (2020). Supporting Privacy, Trust, and Personalization in Online Learning. *International Journal of Artificial Intelligence in Education*. Published.
- Bailenson, J. N. (2021). Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. *Technology, Mind, and Behavior*, 2(1).
- Bedenlier, S., Wunder, I., Gläser-Zikuda, M., Kammerl, R., Kopp, B., Ziegler, A., & Händel, M. (2020). “Generation invisible“. Higher education students’ (non)use of webcams in synchronous online learning. *University of Erlangen-Nuremberg*. Published.
- Braun, V., & Clarke, V. (2012). Thematic analysis. *APA Handbook of Research Methods in Psychology, Vol 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological.*, 57–71.
- Browning, M. H. E. M., Larson, L. R., Sharaievska, I., Rigolon, A., McAnirlin, O., Mullenbach, L., Cloutier, S., Vu, T. M., Thomsen, J., Reigner, N., Metcalf, E. C., D’Antonio, A., Helbich, M., Bratman, G. N., & Alvarez, H. O. (2021). Psychological impacts from COVID-19 among university students: Risk factors across seven states in the United States. *PLOS ONE*, 16(1), e0245327.
- Cao, H., Lee, C. J., Iqbal, S., Czerwinski, M., Wong, P. N. Y., Rintel, S., Hecht, B., Teevan, J., & Yang, L. (2021). Large Scale Analysis of Multitasking Behavior During Remote Meetings. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Published.
- Chen, Z., Cao, H., Deng, Y., Gao, X., Piao, J., Xu, F., Zhang, Y., & Li, Y. (2021). Learning from Home: A Mixed-Methods Analysis of Live Streaming Based Remote Education Experience in Chinese Colleges during the COVID-19 Pandemic. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Published.
- Giannakos, M., Jaccheri, L., & Krogstie, J. (2014). Looking at MOOCs Rapid Growth Through the Lens of Video-Based Learning Research. *International Journal of Emerging Technologies in Learning (IJET)*, 9(1), 35.

- Griggio, C. F., McGrenere, J., & Mackay, W. E. (2019). Customizations and Expression Breakdowns in Ecosystems of Communication Apps. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1–26.
- Han, S., Kim, K., Choi, S., & Sung, M. (2020). Comparing Collaboration Fidelity between VR, MR and Video Conferencing Systems: The Effects of Visual Communication Media Fidelity on Collaboration. *Journal of Universal Computer Science*, 26(8), 972–995.
- Hassenzahl, M. (2010). Experience Design: Technology for All the Right Reasons. *Synthesis Lectures on Human-Centered Informatics*, 3(1), 1–95.
- Hehir, E., Zeller, M., Luckhurst, J., & Chandler, T. (2021). Developing student connectedness under remote learning using digital resources: A systematic review. *Education and Information Technologies*. Published.
- Kehrwald, B. (2008). Understanding social presence in text-based online learning environments. *Distance Education*, 29(1), 89–106.
- Lazar, J., Feng, J. H., & Hochheiser, H. (2017). Chapter 7: Case studies & Chapter 8: Interviews and focus groups. In *Research methods in human-computer interaction* (2nd ed., pp. 153–228). Morgan Kaufmann.
- Leech, B. L. (2002). Asking Questions: Techniques for Semistructured Interviews. *Political Science & Politics*, 35(04), 665–668.
- Mackay, W. E. (1991). Triggers and barriers to customizing software. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Reaching through Technology - CHI '91*, 153–160.
- Marathe, S., & Sundar, S. S. (2011). What drives customization? Control Or Identity? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 781–790.
- Nardi, B. A. (2005). Beyond Bandwidth: Dimensions of Connection in Interpersonal Communication. *Computer Supported Cooperative Work (CSCW)*, 14(2), 91–130.
- Open Broadcaster Software / OBS. (n.d.). Obsproject. Retrieved August 13, 2021, from <https://obsproject.com/>
- Picciano, A. G. (2002). Beyond student perceptions: Issues of interaction, presence, and performance in an online course. *Journal of Asynchronous Learning Networks*, 6(1).

- Rajab, M. H., & Soheib, M. (2021). Privacy Concerns Over the Use of Webcams in Online Medical Education During the COVID-19 Pandemic. *Cureus*. Published.
- Rajpoot, Q. M., & Jensen, C. D. (2015). Video Surveillance: Privacy Issues and Legal Compliance. In *Promoting Social Change and Democracy Through Information Technology* (pp. 69–92). IGI Global.
- Redskaber til online møder*. (2021, June 28). Aarhus Universitet. <https://medarbejdere.au.dk>
- Richardson, J. C., Maeda, Y., Lv, J., & Caskurlu, S. (2017). Social presence in relation to students' satisfaction and learning in the online environment: A meta-analysis. *Computers in Human Behavior*, 71, 402–417.
- Spectralizer*. (n.d.). OBS Forums. Retrieved August 13, 2021, from <https://obsproject.com/forum/resources/spectralizer.861/>
- Stanford University. (2021, March 1). *Four causes for Zoom fatigue and their solutions*. Stanford News. <https://news.stanford.edu/2021/02/23/four-causes-zoom-fatigue-solutions/>
- Stone, C., & Springer, M. (2019). Interactivity, connectedness and 'teacher-presence': Engaging and retaining students online. *Australian Journal of Adult Learning*, 59(2), 146–169.
- Sung, E., & Mayer, R. E. (2012). Five facets of social presence in online distance education. *Computers in Human Behavior*, 28(5), 1738–1747.
- Thompson, S. (2020). Victoria University of Wellington. *A High Fidelity Mixed Reality System for Remote Collaboration*. Published.
- Xsplit Vcam*. (n.d.). XSplit. Retrieved August 13, 2021, from <https://www.xsplit.com/vcam>

Appendix

Appendix A: Setup guide to OBS, link: [https://gitlab.au.dk/au600818/bachelors-project/-/blob/master/Virtuel Webcam Setup Guide.md](https://gitlab.au.dk/au600818/bachelors-project/-/blob/master/Virtuel%20Webcam%20Setup%20Guide.md)