



”Whores, idiots and sub-humans”

A study of the of hate directed at reality profiles and
influencers on Flashback in 2020

Master Thesis in Global Studies

Spring Semester 2021

Author: Stina Lindberg

Supervisor: Michael Schulz

Word Count: 19444

Abstract

Abstract: Hate speech, hateful messages, slander, and other disparaging expression have not been part of the public debate or dialogue but is more apparent on the internet and the online world. A group that has been exposed to a great deal of vitriol on media platforms are influencer and reality profiles. Nevertheless, they are neglected in the research of hostility against public persons. Therefore, this study aimed to examine the hate expressed in the descriptions of reality profiles and influencers, understand what enables users to express hatred, and examine how content differs when directed at men and women. This thesis is based on a single case study of Flashback.se and conducted through content analysis. By investigating 2400 posts in four different discussion threads, the result shows a great coherence with previous research and presents new findings. The findings show that there were 610 statements identified as hateful and coded within eight different categories, representing 365 hateful statements targeting women and 210 hateful statements directed towards men, and 35 against them as a group. The conclusion of the study is that the content is very hostile and hateful, where men are more target to a wider range of hatred due to personal attributes of who they are, and women are primarily exposed to hatred due to how they look, together with sexual harassments. Additionally, in some cases, women are targeted as a group, and the hate against females reflects the social gender norms, misogyny, and gender-based violence in the real world.

Keywords: hate, cyberhate, influencers, reality profiles, misogyny, content analysis

Table of contents

1. INTRODUCTION	5
1.2 AIM AND RESEARCH QUESTION	7
1.3 RELEVANCE FOR GLOBAL STUDIES.....	8
1.4 DELIMITATIONS	9
2. BACKGROUND	10
2.1 FLASHBACK.....	10
2.2 TOPICS AND DISCUSSION THREADS	11
2.3 INFLUENCERS AND REALITY PROFILES	11
3. PREVIOUS RESEARCH AND LITERATURE REVIEW	13
3.1 RELEVANCE TO PREVIOUS STUDY AND RESEARCH GAP.....	15
4. THEORETICAL FRAMEWORK AND KEY CONCEPTS	17
4.1 THE DISINHIBITION THEORY	18
4.2 SOCIAL IDENTITY MODEL OF DEINDIVIDUATION EFFECTS (SIDE)	20
4.3 HATE, EMOTIONS AND FLAMING	20
4.4 GENDER – HOSTILITY ONLINE AND OFFLINE	25
5. METHODOLOGY.....	28
5.1 THE INTERNET AS OBJECT FOR ANALYSIS	28
5.2 RESEARCH DESIGN	29
5.2.1 <i>Case Study</i>	29
5.2.2 <i>Unit for analysis</i>	30
5.3 RESEARCH QUALITY	30
5.3.1 <i>Reliability</i>	30
5.3.2 <i>Validity</i>	31
5.4 SAMPLING AND DATA COLLECTION	31
5.5 CODING	32
5.6 DATA ANALYSIS: CONTENT ANALYSIS	35
5.7 DEFINITION OF “HATE”	35
5.8 METHODOLOGICAL DISCUSSION	36
5.9 ETHICAL CONSIDERATIONS.....	37
6. RESULT.....	39
6.1 THE CONTENT	39

6.2	MEN AND WOMEN	41
6.3	MEN AND WOMEN – EXPLORING THE CONTENT	43
6.4	HATRED DUE TO THE GROUP “REALITY PROFILES AND INFLUENCERS”	48
7.	ANALYSIS	50
7.1	THE NATURE OF HATE IN THE DISCUSSION FORUMS.....	50
7.2	THE NATURE OF HOSTILITY WHEN DIRECTED AT MEN AND WOMEN	52
8.	DISCUSSION AND CONCLUSION	55
	REFERENCES	59

1. Introduction

In the globalized world, technologies have paved the way for communication to take place in a lot of different routes in the digital environment; through e-mails, different apps, file-sharing services, comment fields, discussion forums etc. Besides all the new opportunities this brings, Bladini (2017) points out that there are also various challenges with a more digitized communication world. The traditional boundaries, such as how one conducts conversation and debate, who are allowed in public conversation, have an opportunity to participate and what is being said and mediated, have changed and now include everyone with access to the digital environment. Further, Kaul (2012) means that boundaries related to traditional media such as location, distance and time has disappeared with the digital way of communicating. Fredman (2018) agrees and builds on the argument that the online digital environment can, in many cases, be used for a good purpose. However, it can also be used as a tool to spread hate, threats, radicalization, recruit terrorists, incite physical violence by extremist groups, terrorist networks, incels (involuntary celibate), and conspiracy theorists.

Freedom of Speech is, according to universal consensus, a human right. As MacIntery J declares in the Canadian supreme court of 1986: *“Representative democracy, as we know it today, which is in great part the product of free expression and discussion of varying ideas, depends on its maintenance and protection”* (MacIntery 1986 in Fredman 2018: 307) The right to express one's opinion is central to democracy. In the Nordic countries, freedom of expression exists in the constitutions and has been a right for the people for hundreds of years. The question is if the freedom of expression is equal to expressing opinions and sharing thoughts no matter what the opinion consists of, and if it offends someone (Bladini 2017). Hate speech and hateful messages, slander, and other disparaging expressions have not been part of the public debate or dialogue but are more apparent on the internet and in the online world. This is partly a result of the physical distance between the author of these hateful messages and their audience; also that one can be anonymous or under falsely figured names (Berglind, Pelzer & Kaati 2019 & Bladini 2017). However, anonymity has been considered an essential tool for people to participate and access the democratic discourse. According to the Swedish government, the democratic discourse should be seen *“as an open, common and ongoing exchange of opinions and knowledge in the public space. Its participants can be individuals who participate individually or together with others in, e.g. networks, organizations or parties”* (SOU 2020: 56). Being anonymous can be of great importance to the most vulnerable groups in our society and groups

in less democratic societies to make their voice heard. Unfortunately, anonymity also is used to spread violence, hatred and threatening individuals and groups online (Bladini 2017). Even though the freedom of speech and expression is universally recognized as fundamental rights, and anonymity is an essential tool to the democratic discourse, Fredman (2018) questions if the democratization of speech is paving the way to make censorship less prevalent and create distinct possibilities to combat speech with speech, or if the internet conversely has increased the feasible of speech to cause damage instead?

Consequently, with this as the new state of public conversation, public authorities and other public figures who are vitally affected by the importance of freedom of speech such as journalists and politicians, are being exposed, intimidated and hated in online environments. This is also the case for people who work in more artistic professions, such as musicians, comedians and influencers, who receive hate and threats on a variety of platforms. Shaw (2012) claims that the internet has become a powerful machinery for read, express and disseminate hate, and its millions of different forums provides people who go under anonymous cover to express things they would never utter in person. Further, making hateful expressions and comments has become a sort of amusement for some people, where individuals sometimes create more than just one account to cherish, agree upon and share hateful posts to create a notion of widespread support for the hate and threats. Moreover, the things being said on internet has a far-reaching and determinative impact. Just as Shaw (2012: 280) highlights:

“Today's public consciousness is shaped not in the streets or the parks, but in online editorials and web forums”.

This development has resulted in hate speech online has become everyday life for people outside and inside the public space. It can be found everywhere on the internet; on social platforms, in open as well as closed forums, on websites, in comment fields and so on. Bladini (2017) argues that the hatred and threats are de facto affecting people's way of expressing themselves. Public figures like politicians, journalists, or celebrities fall silent and refrain from communicating, writing, or publishing certain things and values for fear of the hatred and threats that may follow. This means that the consequence of hate speech could be that people delimit themselves and how they act, making democracy and the freedom of speech suffer. In the report “Free to be online?” published by Plan International earlier this year, it is stated that there is far too much consideration is given to freedom of expression on social media and the

web, and no one takes responsibility for the abuse, violence, hate and other types of cyber-bullying taking place in the forums. The report also asks the question “*whose rights are being prioritized and whose voices are silenced?*” (p. 9) This development where freedom of expression seems holds itself captive, as people who have a power to influence, such as public authorities, journalists or other people with large platforms, abstain from participating in public debate or express opinions. However, people inside the political-, journalist-, musician-, comedian-, and research professions have been in the focus of previous research (see chapter 3). Nonetheless, people in the reality-, and influencer profession are a public group that has not been paid attention to in the research.

1.2 Aim and research question

With this as a background, I find it both interesting and important to study these types of environments, what threats and hatred are expressed in order to map them and make the hostility that are perpetrated against public figures in Swedish digital environments visible. Something that has been less studied in previous research is what hatred looks like when it is directed at men respectively women, and if there are differences in the hate and threats being expressed. The overall purpose of this thesis is therefore to make a comparative mapping and analysis from a gender perspective in posts and comments on the largest Swedish social forum Flashback. The gender perspective brings a focus on how online hatred affects women and men, even if other aspects are also made visible. This also includes giving a picture of how hatred and threats can be expressed against each gender. In this thesis, gender should be considered as legal gender.

Consequently, this thesis aims to outline and categorize the hatred and threats conducted against influencers and reality profiles on the internet forum Flashback and examine the possible difference between men and women in the content found in the forum. The central questions are used to further delimit the aim of this thesis are:

- How is hate expressed in the descriptions of reality profiles and influencers in the online discussion threads on Flashback.se?
- How does the nature of content directed against women and men differ in the discussion threads at Flashback.se? By nature, I mean the characteristics, attributes or disposition, such which "characterize" the specific content.

1.3 Relevance for global studies

“The fight for the web is one of the most important causes of our time”

- Tim Berners-Lee (NBC News 2019)

Globalization could be understood as (Ritzer 2011: 2): “a transplanetary process or set of processes involving increasing liquidity and the growing multi-directional flows of people, objects, places and information as well as the structures they encounter and create that are barriers to, or expedite, those flows...”. Moreover, in this day and age, almost everything in society is built on transplanetary networks. The global connection is the key to linking different network systems of communication together. Furthermore, what has become the key player in the network society is the World-Wide-Web. The Internet is one of the most extraordinary global phenomena of modern times, which has drastically changed how we communicate with each other. The Internet is a colossal setting with approximately 4.6 billion people online, of which 3.7 billion use social media. The data located on the Internet is predicted to grow to 44 Zettabytes in the year 2020. Every 60 seconds, about 200 million emails are sent, 4.2 million google searchers and 480,000 tweets are written (NodeGraph 2020).

Although the Internet in many respects stands for good forces, darker forces are constantly present (not to mention deep web and dark web), and even the founder of the World-Wide-Web, Tim Bernes-Lee, understands the concerns for the Internet as no longer working for a good purpose. He identifies three sources of dysfunction that needs to be tackled: there are deliberate and malicious behaviour and use of the Internet, such as hacking sponsored by the states. The spread of misinformation, ad-based models, clickbait and misuse of the user’s data. He states that the world-wide-web's unintended negative consequences are hate speech, polarizing discussions, and bullying in the online environment. Balancing free speech and harmful content is an issue the big social platforms need to deal with every day (NBC News 2019). This is what we call Cyberhate.

As technology and communication are globalized, both hate groups and individual cyber hatred have increased and become part of the online culture. Wherever people are in the world, they are a potential target for threats, hatred and slander through various technological uses, such as chat forums, blogs, social media, etc. However, the Internet's nature is not specific for just the Internet but reflects society as a whole (Perry and Olsson 2009). It could be argued, as

previously mentioned, that globalization of technology and communication has paved the way for a transplanetary flow of hate, breaking former barriers.

1.4 Delimitations

Regarding time and space in this thesis, I chose to examine one location, Flashback.se, and four discussion threads on the site from 2020. Yet, I chose to investigate and annotate a significant amount of comments, 2400, for significant and wide-ranging data material. The data collection material is limited to 2020, primarily because the material is as close in time as possible. There may be opinions that what one wrote on the internet a long time ago has no significance today when, for example, something one says in person is forgotten in another way.

Even though this thesis is delimited to only one location, that is to say; the delimitation will contribute new and exciting knowledge because it is an unexplored theme and group of public persons.

2. Background

This chapter defines the information that may be relevant to the reader to understand the case, such as characteristic features of Flashback, the background to the discussion forum emergence, the discussion threads and how they are moderated.

2.1 Flashback

Flashback has a genuinely long history since its founding (for a review of the years 1983 - now, see Flashback.se for the whole history). The fourteen-year-old Jan Axelsson founded Flashback in 1983. In the first twelve years, Flashback figured as a printed paper and was first launched on the internet in 1995. Another five years later, the platform for discussion "Flashback Forum" received its first post. One of the first implementations a huge link archive, which helped readers find odd and "morally dubious" websites. Freedom of expression has always been an essential feature of Flashback. For example, several controversial websites were published in Sweden the following year. However, these were frequently shut down. The websites were blocked not because they were unconstitutional but because the internet service provider deemed them "morally questionable." As a result, Flashback launched its web hosting service in February 1996. All were welcome to join, but only people who had been blocked from previous servers were able to do so. The web hotel quickly became a thriving meeting place where lots of dissident individuals and organizations gathered.

In 2000, Flashback launched the discussion forum, a site where everyone was welcomed, and freedom of speech was unrestricted. Unlike the Flashback Community, where all discussions were private and only open to participants, this was a forum where everyone could participate. At the time of its inception, the site was known as "Flashback conferences," and it was based on concepts that had arisen due to the Flashback Community's efforts. Over time, the site becomes increasingly popular. However, the service was not developed to accommodate such a large number of users. There was a lack of moderation in certain areas, and the content of the debates started to deteriorate. As a result, in April 2002, a parallel forum called "Flashback Conference Forums" was developed for "Flashback Conferences." In the new platform, the users had to sign up as a member before one could post. Moderators were introduced, and users could also send private messages to other participants. These two forums ran concurrently for

a while, but they were eventually combined into one. Flashback Forum is the name of the forum that exists today (Flashback 2021).

2.2 Topics and discussion threads

There are 16 topics to choose from on the website, ranging from economics, drugs and vehicles to topics such as politics, lifestyle or culture, with several sub-topics. In all topics, one can create or write in existing discussion threads if one is a site member. All threads can be read, even if one is not a member. For every sub-topic, there are moderators which ensures that the discussion rules and quality are followed. If a thread or a user breaks the rules, threads can be locked or thrown in the trash, allowing them to read only the threads without commenting. Users often get warnings, but in some cases, they are blocked from further postings.

The choice of Flaschback.se is primarily due to the great emphasis on freedom of expression, the size of the web page, the moderation (for example, that not everyone can go into threads and delete posts), the activity on the page and the thousands of threads that exist. Further, all post can be read by anyone, and the users are anonymous.

2.3 Influencers and reality profiles

We've seen social media's popularity skyrocket in the last decade. According to the We Are Social survey from January 2019, 3.484 billion people use social media, accounting for 45 percent of the world's population. Inevitably, these individuals turn to social media influencers for guidance in making decisions. In social media, influencers have developed a reputation for their insight and experience on a particular subject. They create vast followings of passionate, committed people who pay close attention to their opinions by making frequent posts on the topic on their favourite social media platforms (Influencer Marketing Hub 2021).

There are many types of different influencers, all from professors, teachers to celebrities and bloggers. Many people who participate in reality shows in Sweden start working with their social media after their participation and get a lot of followers before-, during- and after the show. Some people work with social media full-time, for example, by cooperating with companies or promoting different lifestyles. Mostly, the people are young, doing their first appearance for a large audience, meanwhile others participate in the reality shows for the second, third or even fourth time, in some cases, to get more followers. Influencers and reality

profiles are all usually placed in the same compartment when discussed; hence I use both terms. Familiar to all the people mentioned in the collected material is that they are in some way very active on their social media, have or currently work in collaborations with companies and have a lot of followers.

3. Previous research and Literature review

This chapter presents an overview of previous research in the field or related fields. National, as well as international reports, articles and relevant publications will be provided. Hatred, threats and violence are world phenomena found in an incomprehensible amount of research and often recur in psychology, humanitarian, social and political science, economics, gender studies, peace and conflict and many others. As of the study of human behaviour, hatred and violence play a key role in understanding how we interact and communicate. Furthermore, with the internet's success and scope, the way people online mutually connect, the exploration of cyberhate, online hate and hate speech, and humans' vulnerability and activity over the internet has grown in the last years.

A considerable amount of the research is quantitative data collection through surveys, machine retrieval from web forums or an exploration in official documents and guidelines on how to work against the problem. In the report "*Hat och hot på nätet: En kartläggning av den rättsliga regleringen i Norden från ett jämställdhetsperspektiv*" Bladini (2017) presents an overview on how Nordic countries operate when someone is exposed to hatred, threats, and other violations online. There is widespread confusion about working against online hate and uncertainty about what is criminal and how to handle these types of violations. This exists among the affected individuals and groups and judicial authorities, resulting in challenges for the rule of law. Moreover, there is a prevalent distrust of police work regarding cyber hatred and violations, where the population consider authorities to overlook online-based hatred and harassment.

Further, several studies demonstrate that those who are particularly exposed to online abuse are people who take part in the public discourse. Commonly these people have employment such as politicians, researchers, artists, authors or journalists, and often women (Journalistförbundet 2014; Barton & Storm 2014; Eggebø. & Stubberud: 2016; Eriksson Almgren., Ekblad, & Sjören 2017; Totalförsvarets forskningsinstitut, hence FOI, 2017, FOI 2020). Gardiner (2018) published an article with findings that women, mainly black, Asian, and ethnic minority journalists, received more hate, despite the topic of the article they publish, are subject to abuse in a broader range than their male colleagues. This is in line with research findings that other minorities are especially hated, as people of colour (Kang 2000; Pantti et al. 2019), or for example with people with an Islamist religious belief (Awan 2014, Awan 2016) are targets for

a large scale of hateful discussion online. Further, Baider (2018) shows that people with non-heterosexual identity are targets for hateful narratives in the online world.

There is a joined vision of the online environment paving the way for movements to elaborate on a global scale (for example, see Kellner, 2003; Kahn and Kellner 2004; Hara, N and Huang, B. 2011; Maireder, A and Schwarzenegger, C. 2012; Mattoni and Treré 2014; Del Vicario et al 2016), for instance, to spread information, organize meetings and demonstrations and other activities. However, the new media has set the stage for campaigns focused on extremism and radicalization. White supremacist groups, far-right wings, and other extremist groups have been extensively studied in this field. Kahn and Kellner (2004: 94) highlights ‘the internet is a contested terrain, used by Left, Right, and Center of both dominant cultures and subcultures in order to promote their agendas and interests. The report “Digital Slagfält” made by the Swedish Defence Research Agency on behalf of the Swedish Government provides an insight into how people demonstrate radical national ideologies on the internet.

Moreover, studies have focused on victims' experience of cyber hatred. Studies show that the hatred that previously took place in school and activities has transferred to the internet and turned into cyberbullying. It has become prevalent for young people to receive cyberbullying, which usually results in anxiety, mental health issues, insecurity, worry (Tynes 2005; Staude-Müller, Hansen & Voss 2012; Mascheroni & Ólafsson 2014; Selkie, Kota, Chan & Moreno 2015; Fridh 2018; Blaya 2019) and a higher risk of suicide (Hinduja & Patchin 2010). Studies have also shown that victims expose others to hatred and violence (Leung et al., 2018). Besides, there seems to be a correlation between being exposed to hate and radicalization and extremism (Tynes 2005; Foxman & Wolf 2013), along with an absence of trust of people, greater discrimination, increase of prejudice and hate crimes outside of the internet (Näsi et al., 2015; Keipi et al., 2017).

Gender studies have in recent years become a growing field in this research area. Women’s Media Center (2021) points out that there is a unique gender-based abuse online attacking females. This online harassment of women could also be known as Cyber misogyny or Cybersexism. Scholars have pointed out that misogyny culture is prevalent in the online world (Barak 2005; Bartow 2009; Citron 2009; Ritter 2009; Franks 2012), even though women account for a larger part of the social media users (Heil and Piskorski 2009; ConnectAmericas 2021). In the study ‘ *violence against women: an EU-wide survey*, Fundamental Rights Agency

explored online stalking through text messages, e-mails and online messages through an extensive collection of data. Young women suffer the most from cyberstalking, where almost 1,5 million women between the years of 18 and 29 had been targets of online stalking the last 12 months (4%) compared to women 60+, where 0,3% had experience. Further, 11% of the women had received offensive, unwanted, sexually explicit messages and inappropriate sexual suggestions on social media forums.

As the previous literature have explored, the online hate environment often targets people taking part in the public discourse, such as journalists, politicians, researchers, artists or authors. Women in these professions are particularly exposed to hatred and threats online. A newly produced research study by FOI (2020), examines the nature of digital hatred on Flashback and describes how women in these professions are particularly exposed. In addition to journalists, comedians, politicians, musicians and influencers were also examined, and the results showed that they were overrepresented in several categories, such as sexual harassments regarding the nature of hatred. FOI mapped digital hatred and how it differs for different gender categories and different occupational groups. The analysis show that the hatred targeting women and men look different. Women are subjected to more sexual harassment and appearance-related insults, while men are more often insulted because of their profession and skills. Of the women included in the survey, 55 per cent were exposed to hatred, while the corresponding hate for men was 41 per cent. The study also shows that the most significant proportion of people exposed to hatred are female journalists, where 67 per cent of the people are exposed to hatred, followed by female influencers of 63 per cent. In that study, a codebook based on the material was used, which will also be the inspiration for this study.

3.1 Relevance to previous study and research gap

However, influencers and reality profiles are a group of people within the public discourse who are not yet studied. There is a rise of the “influencer/influencer marketing” and reality profile professions, where people are dominating the social media platforms, such as Instagram, Twitter, Facebook, Tik Tok etc. Influencers and reality profiles often work with their social media platforms and podcasts, commonly with different collaborations. These profiles have a huge audience, often younger people, yet adults and older people. The platform could be used, for instance, to influence, spread messages, knowledge, pursue political agendas or work for particular campaigns. Yet, their comment fields is overcrowded by love and hate, negativity

and positivity, and in some cases, the hate has resulted in real tragic events. Two people in the last year, a Japanese tv-profile and a polish influencer/model took their own life, and there is evidence of cyberbullying contributing to the deaths (Reuters 2020, Euro Weekly News 2021). This has brought up the question of cyberbullying towards influencers and reality stars, and all over Instagram, Swedish and international influencers and reality stars shed light on the problem of cyberhate. Influencers and reality profiles also lack senior employees to turn to, no safety representative or one who cares for the work climate, or a union that can help in unfair conditions. Even if they can moderate their comment field, they cannot impact discussion forums, such as Flashback. With this as a background, and with the group of influencers as a forgotten public group in the previous research, this thesis intends to contribute to the research gap and focus on this group.

4. Theoretical framework and key concepts

The classical theories regarding communication or freedom of expression were developed when social media platforms and current communication mechanisms did not exist. As the literature review shows, the research examines the phenomenon behind cyberhate, for example, the victim's experience or the people behind the hate, instead of how and against which net hatred is expressed. Bryman (2012) explains when a researcher wants to study a phenomenon where none or a small amount of research has been done in the past, the researcher could take a more exploratory stance. Rather than testing a theory, because of the still less developed theoretical framework of the nature of hatred online, this thesis will take an exploratory standpoint leaning on grounded-theory model, to a better understanding of the existing problem. Content analysis is suitable for data that require some degree of interpretation. In any case, this essay is inspired by grounded theory as it aims to contribute to the theory by testing and coding the material.

These theoretical points intend to thus help understand, explain and evaluate what has emerged and further it to a more principled level. Due to this relatively new research field, we therefore have few relevant theories that have been developed and could be applied to the research problem at focus of this study. I do, however present theories that can be seen as theoretical frameworks that not directly can be applied or “tested” but contextualize and understand the focus of the research problem. The function of theory in this essay is to frame the problem, make sense of it and present different angles of approach to this problem. The theories in chapter 4.1 will work as framing the problem and context in which hate speeches exist. Perspectives of deindividuation, anonymity and group identity are useful for analysing how people express themselves when they are anonymous or belong to a group. The experience of the disinhibition effect and the SIDE model can as well exemplify the background to the hatred found in the material and why people do not express hatred in the same way eye to eye as they do behind a computer screen. Thus, this theory describes the precondition for hostile usage of the internet and ideas that enable or make it feasible for people in the discussion threads to be hostile and express hatred. However, this cannot be examined and measured in this thesis, but it provides insight and frame in which context hate speeches functions. The theoretical concepts presented in chapter 4.2 aims to explain how the hate and insults are founded, who they are directed to and how they can be articulated in various ways. Hate takes a wide range of expressions and people experience it in different ways, however this theory can clarify the content I have been looking for in the material. The third section of the theoretical framework

aims to focus on the gender aspect, which is addressed in the third research question. The fact that hate research seems to have discovered that there are differences between men and women, and therefore this section will focus mainly on men and women. This section of the theory covers both online and offline harassment and aggression.

4.1 The Disinhibition theory

According to Suler (2004), researchers, individuals and users of the internet have noticed how people say, perform, or engage in things on the internet that they would not conventionally say and do in the up close and personal world. The internet is a space where the consumers and producers of material and texts work within a more unwind environment, where users feel less controlled or restrained, resulting in the communication becoming more straightforwardly. Although anonymity may seem like a tool for sharing inner feelings, fears, desires to seek help from others, we observe a much darker side to which people use their anonymity. Anger, hate, threats, violence, harsh criticism, slander, different types of illegal or offensive pornography, crime, abuse and other types of harmful and adverse content are as well a significant part of the content on the internet. This is a sphere where the visitors would not visit in the real world and Suler (2004) calls this the *toxic disinhibition*.

Whether a more healthy environment or toxic, Suler (2004) identifies six factors involved in creating the disinhibition effect. Firstly, there is dissociative anonymity, where a person can use a real, false or stolen identity. The users online can remain somewhat anonymous, and when this opportunity appears, the actions online are separated from the real identity and lifestyle and contributes to a notion of feeling less vulnerable about self-disclosing and acting out and cannot be related to the behaviours and sayings on the internet for the rest of their lives. In this cycle of separation, users do not need to be responsible for their actions by recognizing it inside the entire set of an incorporated on the web/disconnected character. The online self turns into a compartmentalized self. On account of communicated threats or other degenerate activities, the individual can turn away duty regarding those practices, as though superego limitations and good intellectual cycles have been incidentally suspended from the online self and could convince themselves that the actions online are completely detached from the internet-character. Secondly, the notion of being invisible, which somehow overlaps with anonymity, comes into play in the online environment. In chat rooms or discussion forums, other users may not know who is present, and in these types of sites, there is likely to be hundreds and thousands

of users. The feeling of being invisible also allows people to visit territory and do things they would not otherwise dare to go. The ability to shield oneself from others' body language, facial expressions or other expressions also causes people to act differently than they would in reality.

Thirdly, the asynchronicity in cyberspace, where the immediate reaction or answer to one's message disinhibits users. The absence of direct feedback and responses strengthens a few practices and stifles others and effectively shapes the continuous progression of self-disclosure and conduct articulation, typically toward adjusting to normal practices. The asynchronicity could also be used as a "running away" tool or "leave behind" after posting hostile content.

Fourthly, the users experience the content "as a voice within one's head, as if that person's psychological presence and influence have been assimilated or introjected into one's psyche" (Suler 2004: 323). This is referred to as solipsistic introjection, where the other user and oneself become a character in mind. Just as fantasizing about different occasions and situations within one's head, a person feels free to express however they want because the feeling is that it only plays out in their mind. Further, the fifth factor is a dissociative imagination, where the person thinks of it like "it's only a game" where the characters created online are working in a different world. Once the users turn the computer off and starting to be a citizen in the real world again, they feel like they have left the game and their game-character online. What happens online has nothing to do with reality.

The last factor Suler (2004) identifies is the minimization of status and authority. In a face-to-face meeting, elements such as power, status and authority come into play in the way people dress, express themselves and other personal settings. However, in the online meeting, users cannot identify these factors as easy, resulting in a reduction of authority and status in cyberspace, and even if the authority figure is known, their presence, status and power have less effect online. This is an opportunity for everyone to voice themselves with an outcome of a more democratic internet. Communication skills are a real asset for creating one's own authority and status online, together with persistence, the quality of one's ideas, and technical knowledge. Further, the absence of fear of being punished or criticism from authorities tends to make people willing to speak out and misbehave more.

4.2 Social Identity Model of Deindividuation Effects (SIDE)

The SIDE-model implies the effect and use of anonymity in Internet-based communication and how the individual in these contexts tends to de-individualise and consider other individuals on the Internet as de-individualised (Lea et al. 2001). The notion of de-individualisation and anonymity is essential for this thesis because the effect of de-individualisation and anonymity in combination with group norms may underlie or affect how one expresses themselves, in this case, expressing cyberhate. Lea et al. (2001) describe the SIDE model as based on two different sides: the cognitive side and the strategic side. The cognitive side deals with the perceived social identity. On the Internet where anonymity dominates, a forum member's identity becomes more prominent than the personal identity. Following group norms created on the Internet becomes more prominent, as people identify more strongly with the group than with their real-world identity. The strategic side is about the individual's strategy to use anonymity for personal gain. The individual uses anonymity to present his or her opinions and positions in the way he or she wishes. Additionally, anonymity enables individuals to express opinions that non-anonymous and public contexts would have been more lenient. Christopherson (2007) suggests that a combination of the two sides tends to increase provocative expression as well as make them cruder.

4.3 Hate, emotions and flaming

“[...] the term hate speech is understood as any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor”.

- Definition by United nations in STRATEGY AND PLAN OF ACTION ON HATE SPEECH (2019)

Hate as an emotion is explained by Ahmed (2014) as "a negative attachment to another that one wishes to expel, an attachment that is sustained through the expulsion of the other from bodily and social proximity" (Ahmed 2014: 55). Further, when hate becomes a crime, it is often because of an individual's group identity, such as religion, sexuality, gender or religion. In the case of hate crimes, the violence against one individual works as an act of violence against the whole group. An act of hate is described as a hit in the emotional place where it hurts the most. Matsuda (1993: 24 in Ahmed 2014: 58) writes, "The negative effects of hate messages are real

and immediate for the victims. Victims of vicious hate propaganda experience physiological symptoms and emotional distress, ranging from fear in the gut to rapid pulse rate and difficulty in breathing, nightmares, posttraumatic stress disorder, hypertension, psychosis and suicide". A hateful or insulting word does not need to appear directly as hateful, for example, the word "Paki". However, some words have the historical context of being used in a way to discriminate, dehumanize or insult a group. The nature of hostile language allows us partly to understand if the speech "works" or "fails". However, there is a need to ask how it affects the bodies that become targets for hate (Ahmed 2014). This question would be more suitable for future studies to understand how it affects the people exposed to hatred.

In an extensive review of classic and more modern conceptualizations of hatred, Royzman, McCauley, and Rozin (2005) identify hatred as one of the most harmful affective phenomena in humankind's chronicle. The notion of the destructive consequences of hatred in people's lives, Fisher et al. (2018) reports, are extensively documented in previous research. Hate has been outlined in several ways in the publications, all from being an emotional attitude or a simple emotion, a syndrome, a form of generalized or generalized assessment, a motive to degrade people or as a normative judgement. Scholars that have study "hate" acknowledge hate as a powerful, negative emotion or emotional phenomenon and evolve in situations of different forms of abuse and degradation against someone (Sternberg 2003; Royzman, McCauley, and Rozin 2005; Aumer-Ryan and Hatfield 2007). Meanwhile, other scholars do not define hate as a feeling. Nonetheless, hate is considered to share some of the features of feelings of intense dislikes, such as moral distaste, contempt or various forms of anger (Frijda 1986; Fitness and Fletcher 1993; Oatley and Jenkins 1996; Halperin 2008).

"Hate speech" is a term, most used by the academic world, government officials, and the public to outline harmful content in the online world. However, Parekh (2012) states that it appears to a variety of forms of speech could be defined as hate speech, just as well as that they cannot fit into the definition. This relies a lot on the context in which the speech is made, which in itself may or may not fit into the definition of hate speech. The restricted forms of speech are not in all cases affiliated. The speech could be abusive and insulting, but not perceived threatening, and speech could also communicate or advocate different views, without requesting for action. There could be an expression of dislike of a group, however not expressing hatred. One could express resentment towards a group, but perhaps not outright hatred, while the expressions can be so subtle that they are not offensive or insulting to anyone who takes part in the speech.

Additionally, speech could bring a critical or calumnious perspective of a group, but with no harm aspirations and take a patronizingly indulgent attitude. However, some forms of speech can be identified as immoral, unacceptable, or offensive without being allegedly similar to hate speech; therefore, all hateful speech should not all be grouped together and addressed in the same way. Parekh (2012: 40) writes: "Hate speech is a distinct kind of speech, and much conceptual confusion is created – and the net of prohibited speech unduly widened – by subsuming all forms of uncivil and hurtful speech under it. We may therefore rightly proscribe hate speech without also proscribing insulting, abusive, denigrating, and similarly objectionable speech."

Nonetheless, Parekh (2012) declares that even though hate speech is not equal to abusive, degrading or insulting speech, we still have reasons to prevent such speech. Because the term itself stresses that one hate something or someone or is filled with hatred, one of humankind's strongest emotions, the term itself is inadequate. Regardless, the term is widely used, and there is no clear alternative; it still used to capture a considerable amount of different hateful, degrading, abusive or insulting speech. Siegel (2020) agrees with Parekh (2012) and presents arguments to emphasize that a definition of hate speech may not be beneficial as the variation in the content might be comprehensive and depending on the context. It can be problematic to identify what is hateful when the author of the post uses different epithets to describe a person, while it is easier to identify insults and slurs. Moreover, Siegel (2020) contends that it can be highly comprehensive or limited due to the different notions of hate speech. As Parekh (2012), more comprehensive definitions recognize a wider scale of speech directed against a particular or easily recognizable person or group based on arbitrary or normatively irrelevant characteristics.

KhosraviNik and Esposito (2018) outline online hate as communication and practice. They point to Kopytowska et al. (2017) findings of how the interactive and intertextual nature of the internet and online environments provides a space for groups and individuals to manifest their opinions, often including hateful ideas. Cyberspace also offers other users to engage in these forums, resulting in a discursive spiral of hate. Further, the authors explain that the global discussion of cyberhate has to a great extent been visualized around the ideas of "trolling" and "flaming", utilized as umbrella terms for numerous online harmful practices identified with uninhibited online communications. Early definitions of flaming refer to that people express themselves harsher online and that the opinions are reasonably solid and provocative. However,

flaming is at present characterized by "profanity, insults, negative affect, and 'typographic energy' such as capital letters and exclamation marks" (Jane 2015: 66) and that flaming is hostile, aggressive, and offensive or expletive language. In the shift in digital communication, flaming can also be perceived as a natural component, where affective expressionism stands over logical resonance. KhosraviNik and Esposito (2018) refer to the term trolling on the other hand, as overall insults and sick jokes, to more extreme threats of violence (sexual, e.g. rape, physical, e.g. murder or assault). However, a characteristic of trolling is that the users who are using trolling often tricks people into natural or positive discussions and then shifting the nature of the discussion into a more aggressive environment. Together, trolling and flaming is creating a hostile environment in cyberspace.

Even though there is widespread use of both terms, KhosraviNik and Esposito (2018) point to the absence of principled definitions and univocal terminology. Terminology is a vital part of the conceptual, methodological, and epistemological challenge of cyberhate. The first wave of cyberhate research arose in the very moment of the arising of chat forums; however, the polarized debate about computer-mediated communication (CMC) seemed to strand the studies. The "second wave" instead focused on definitions and theoretical models. However, the consistent change in the online world and its platforms makes the application of the early theories prove to be practically, epistemically and conceptually difficult, and a new way of practising the digital environment through new devices also increases the influence of cyberhate. Cyberhate is not just an isolated online dynamic anymore, but a socially pertinent digressive practice (KhosraviNik and Esposito: 2018).

According to KhosraviNik and Esposito (2018), the last years of conducting research in this area has questioned existing theories, methods and terminologies related to online hate. The case of online hate is very complex and hard to grasp in its entirety, making researchers overlooking the topic and avoiding it. Some scholars have suggested that flaming needs be extinguished and instead develop a more ethnographic approach to online hostility (Lange 2006 in KhosraviNik and Esposito 2018). In the development of critical research on online hostility, some matters need to be considered. Scholars need to keep away from taking part in an ethical classification that favours one side. However, the study may have a chance crashing in an ethical relativism and academic separation, "which ignores the tangible social impact of online hate and may result in an underestimation and trivialization of the phenomenon" (KhosraviNik and Esposito 2018: 51).

Women's media center (2021) refers to internet hostility as online abuse. Its taxonomy consists of a wide range of tactics and behaviours online, ranging from embarrassing persons to death threats. Some tactics involve predators and grooming, cybermobs, google bombings to *DOS*, corresponding to the term "denial-of-service" which refers to an intrusion that prevents users from accessing a website or network resource tactics through direct contact harm a victim, which are not vital for this thesis. However, some tactics listed are prevalent for the research questions:

- *Cyber-exploitation, Nonconsensual Photography or "Revenge Porn"*. Without the consent of the recipient of the photographs, sexually explicit images are distributed.
- *Deadnaming*, a form of personal harassment in which a target's previous identity is exposed against their will in order to cause them harm. This method is most widely used to identify members of the LGBTQIA group who have changed their birth names for a number of reasons, including professional prejudice or physical risk.
- *False accusations of blasphemy*. Women face online threats all over the world, but they are especially vulnerable in conservative religious countries, where blasphemy is illegal and honour killings are common.
- *Flaming*. Threats, insults, curses, and profanity as earlier described.
- *Gender-based Slurs and Harassment*. The use of language, threats, profanity, and, in some cases, photographs to express animosity towards girls and women because they are women is known as gender harassment. Harassers sometimes use terms like "bitch," "slut," "whore," or "cunt," as well as comments on women's sexual looks.
- *Sexual Objectification*. Attackers also objectify their victims, using distorted images and sexually suggestive representations of their bodies as examples.
- *Slut shaming*. A gender-based harassment shaming girls for their sex-life.
- *Threats*. Often including, rape and death threats.

Cyberbullying is an additional sort of practice or tactic of cyberhate or cyber abuse. Kansara and Shekokar (2015) address this as the internet has given the bully a new place of harassing people for a bigger audience to see. Cyberbullying consists of insulting, threatening, embarrassing or other types of harassments of people. It can take both a textual form, but also through recordings or images. KhosraviNik and Esposito (2018) address the cyber world's notion to be more playful and dynamic; it could be problematic to investigate such a severe

phenomenon as cyberhate in the context. There have been some difficulties when framing cyberhate within a social, material, and ethically relevant narrative about criminal aggression. Equally, this issue has hampered the recognition of online animosity as a real strategy to suppress and influence marginalized groups who might gain more from the cybersphere's participatory capacity, especially those who have traditionally questioned society's unequal distribution of power.

The normalization of online abuse as an entire act of digital citizenship has been aided by the assignment of a childish and goliardic nature to the phenomenon of trolling, compounded by remarks like "don't feed the trolls," "it's just boys being boys," and "it's just the Internet." (KhosraviNik and Esposito 2018). In more recent research, "flaming" together with "trolling" are called cyberhate and hate speech instead. They think of these terms as "a viable starting point to foster a more mindful approach to the violent and exclusionary nature of online hostility" (KhosraviNik and Esposito 2018: 51).

4.4 Gender – hostility online and offline

The concept of gender has lots of angles and does not always have concern the differences between men and women, but can also focus, for example, on social or cultural distinctions. Since this study intends to examine this very difference, in addition to that these types of discussions often targets hostility towards females or males because of their sex, and that the targeted people in many cases come from relatively similar backgrounds, the predominant perspective in this study will concern women and men with the legal sex.

The European Commission defines gender-based violence (GBV) as "directed against a person because of that person's gender or violence that affects persons of a particular gender disproportionately" (European Commission on GBV). GBV can take different forms of physical, sexual or psychological. Even if these are often referred to as physical acts, GBV is a critical factor of online violence. Online violence could also be linked to violence in real life or limited to the online arena. It could include hostile acts such as threats, unwanted, offensive or sexually explicit statements, sharing intimate images or videos, incitement of violence, stalking, harassments etc. (European Commission on GBV, June 2020). However, this does not mean that women receive more online violence than men. Nadim's and Fladmoe's (2019) findings show that men experience more online harassment than women. Further, the findings show that men

receive comments about their opinions and attitudes; meanwhile, women's harassment aimed at group traits affects both men and women equally. On the other hand, women who are targets of online harassment are more likely to become more cautious in expressing their opinions publicly than men. The author also discovered that when harassment becomes more violent and geared toward group characteristics, gender disparities become more pronounced.

The "Council of Europe - Gender Equality Strategy" (2017) describes hate speech as a phenomenon found everywhere online as well as offline: in social circles, in the public space, at work or in school, in the family relations, social media etc. Sexist speech is a massive part of hate speech, and even if this kind of hate speech has taken new forms and are communicated through various channels, it still is based on the principles of the persistent unequal power relations between women and men. This kind of speech is a form of violence against females, a form of violence that reinforces and aggravates discrimination against women.

Additionally, feminist scholars focus on misogynistic violence. Jane (2020) draws on the growing body of literature on cyber hate directed at women that shows that on the contrary to claims by commentators in a range of context that it's just words on a computer screen, the suffering from this new hostility is embodied, tangible and in all matters real. It causes not only psychological harm but reputational, economic, social and political harm. Also, this contributes to the existing gender-related digital divides. She claims that this cyber hate directed at women is based on far older traditions. Jane (2020) suggests that this demonstrates misogyny's perseverance, the long-term effects of structural gender inequity, the difficulty of social issues arising from machine-human experiences, and the importance of feminist advocacy. She describes some characteristics of the cyberhate directed against women. It includes profanity, aggressive and sexualized imagery, overt, ad hominem invective, and the framing of coercive sex as an all-purpose correction are all signs of the discourse. Further, there are also questions about intelligence, mental health and how attractive women are, often referred to as sexual attractiveness, where women's "fuckability" or "rapeability" is a matter of discussion.

However, regardless of recognizing misogynistic violence as a social problem, scholars have dismissed or neglected it as a form of gender-based hate speech. The definition of hate speech by The Organization for Security and Co-operation in Europe (OSCE) follows: the expression of hatred towards an individual or group of individuals based on "protected characteristics", such as "membership to some specific social group that could, on its own, trigger

discrimination" (OSCE 2009: 37– 46). The classification of these characteristics and whether or not they include gender is still up for debate. Sex, gender identity, and sexual orientation are all factors that have been left out of the equation when definitions or campaigns have been implemented. Misogyny and homophobia, for example, do not feature in the web glossary of words created as part of the same movement. There is an omission of gender from most systemic concepts of hate speech, which results in a blind spot. This partly means that there is an institutional inability to recognize gender as a social dimension that, in and of itself, is sufficient to cause hate speech. The widespread notion of gender equality has been partly achieved, leaving only a legacy of misogyny in popular culture, which may explain the aversion to discussing sexist speech in terms of hate speech. Hate speech is assumed only to be targeting people of vulnerable minorities (KhosraviNik and Esposito 2018).

Further, KhosraviNik and Esposito (2018) describe how feminist scholars recognize that online and offline misogyny have established a clear connection between online abuse and the gendered nature of interactive environments, while it works as a mechanism for maintaining patriarchal social hierarchies. Feminist studies acknowledge for a full congratulate of misogyny and sexist speech as a form of hate speech. KhosraviNik and Esposito (2018) argue that women's precarity in the cybersphere is based upon the users' conformity to gendered social norms. Men and women are populating the cybersphere in different ways, where women should be known to inhabiting the cybersphere as precarious subjects. This is a result due to the assumption of the premise that adhering to these social norms results in a "differential distribution of recognizability" (KhosraviNik and Esposito 2018: 53) and that any decision to violate them results in being differently subjected to harm and crime. Additionally, when women participate in the online public sphere, it could quickly result in a violation of gender ideology's social norms, triggering harmful, sexualized speech to restore "order". Defamatory speech against women is primarily responsible for establishing and maintaining women as precarious objects in the cybersphere. This is accomplished by applying precise illocutionary, and perlocutionary control to eradicate those groups, legitimise discrimination against them, advocate violence and hatred, and trigger shifts in attitudes and behaviours (KhosraviNik and Esposito 2018).

5. Methodology

This chapter elaborates on the research design, samples and collection of data, and the following analysis method. This thesis aims to outline and categorize the hatred and threats conducted against influencers and reality profiles on the internet forum Flashback and examine the possible difference between men and women in the threats and hatred found in the forum. To achieve the aim, this thesis will be carried out in the form of a case study, where the phenomenon of hate and threats will only be examined in the social forum "Flashback". To answer the study's questions and achieve the purpose, content analysis will be used as a research technique. A qualitative methodological approach as well as a quantitative approach were implemented for the analysis of the data to answer the research questions. The quantitative analysis is used for the purpose of an understanding of the amount of hate the influencers and reality profiles receive, and the distribution between different categories of derogatory and abusive speech. The qualitative analysis is used for an understanding of what the content de facto express, and whether the content, the way of expressing hatred and the statements itself differ when directed at the respective gender.

5.1 The internet as object for analysis

The Internet has offered many new research possibilities for social research, and its significant amount of data is material for quantitative as well as qualitative content analysis. The two main e-research areas existing for research identified in Bryman (2012) is;

1. the World-Wide-Web sites and pages and
2. using the World-Wide-Web and online communications as a medium for collecting data, either from organizations or individuals.

Although these two online study fields' choice and classification are somewhat arbitrary and appear to shade one another, these areas provide a foundation for a quite extensive outline of the field. Bryman (2012) refers to four main challenges that are worth considering. Firstly, the researcher is required to find websites relating to the research question(s), and secondly, the researcher needs to choose very applicable keywords for searching online. However, the search engine could be affected by several factors, and the different search engine could bring different results. The researcher should therefore consider using different research engines. Be as it may, this thesis uses a case study - the internet forum Flashback, and will not be affected by this issue.

Further, web pages are not permanent creations but are in a continually changing form, and new web pages are created all the time, while others are taken down, re-created or changed - in terms of appearance or content. Therefore, researchers examinations and analysis could be of content that no longer exists, are updated or changed (Bryman 2012). This factor is essential to this thesis, where users can delete their posts, and threads are constantly moderated and can be locked. If threads are of poor quality based on Flashback's guidelines, the thread may end up in the Trash.

Bryman (2012) also states that further issues need to be handled when investigating online communities and interactions, such as posts in a discussion forum and chatroom interactions. As in this thesis, the context will be to observe and analyse the statements without any participation. Critics may claim that it is slightly ethically dubious about observing without declaring one's participation (see chapter 4.7 for ethical discussions). The researcher must decide on his or her role as an ethnographer and in what way one should participate. Internet and the digital environment have offered researchers invisible while observing and interacting within the context of the research. One of the highlights of conducting examination through online discussion forums is that researchers may decide to stay imperceptible among members by noticing others' cooperation without self-inclusion (Hine 2015). For this study, I will not participate or get involved in any discussion, just be observing. Since I am investigating a notably anonymous environment, where the language can be very raw, I do not want to disrupt or affect the users' communication. If the users knew of the examination, they could be changing the way of authoring the statements. Additionally, the posts were posted before starting the data collection.

5.2 Research design

5.2.1 Case Study

A case study is concentrated and detailed and is often considered a study to examine the specific cases' complexity and uniqueness. The case is often related or linked to a particular location (Bryman 2012: 67). Further, a case study is often a present-day in-reality problem concerning other comparable cases and historical and social contexts. A character of case studies is that there is a focus on a specific unit of analysis (Darian-Smith & McCarty 2017).

The subject of this case study and location is the internet forum Flashback, one of the most visited internet forums in Sweden, and the object that will be studied is the insults, hate and slander statements posted. As Darian-Smith and McCarty (2017) explain, a case study can be both qualitative and quantitative, or a mixture of them both, and often commence with a description of empirical observations of the subject over a significant period of time. Further, the authors recognize that it is necessary when conducting a case study to understand the historical dimension of the specific case. As mentioned in chapter 2, the website has a historical character of protecting the freedom of expression and anonymity of its users, which possibly contributes to the way users express themselves and which topics are discussed. Further, I chose to study Flashback because on other social media, where the comment fields are personal or linked to, for example, a company, they are usually moderated quickly, more consequently and stricter than the threads on Flashback.

5.2.2 Unit for analysis

In content analysis, the researcher consider different “unit of analysis” to answer the research questions. For example, it could be actors, words, subject or themes. By coding due to subject and themes, the researcher could categorize the phenomenon of interest (Bryman 2012). This thesis sets out to examine the characteristics of hate directed to reality profiles and influencers, together with investigating if it differs when directed to men and women. Therefore, the choice of coding into subjects and themes will be the best choice for this thesis. The units that will be the posts and statement posted in the chosen threads at Flashback.se.

5.3 Research quality

5.3.1 Reliability

Bryman (2012) present reliability as a matter of whether the study is repeatable or not: that is, if another researcher would have a similar result. Reliability refers to if the concepts in the study are consistent and if the measure is persistent. One challenge for reliability is the annotations of data in content analysis, which the researchers bias could influence. Bolognesi, Pilgram and van der Heerik (2017) argue about three considerations for content analysis. Firstly, there is stability, meaning that the data annotations would not change if the analyst annotated the same data in a second round. Secondly, if another researcher would annotate the same data, addressed as replicability, would the annotations remain the same? Lastly, accuracy, whether to what extent the annotation process conforms to its specifications and provides what it is intended to

provide. However, the researcher should not choose to simplify the coding scheme in order to achieve high reliability. The analyst needs to balance highly replicable and highly accurate coding schemes (Bolognesi, Pilgram and van der Heerik 2017).

To achieve a high reliability for this study, the coding scheme will consist of the code, a definition, and examples of words or phrases from the material so other analysts can use the codebook in the same way. Definitions and examples will also make the annotations as safe as possible and that the annotation would take place in the same way if it needed to be re-applied. Further, definitions and examples will help to achieve a high-quality of correctness.

5.3.2 Validity

Validity can be described in a variety of ways. The validity of the research refers to the accuracy of which samples are taken or the degree to which a test accurately tests what you want to know. Further, external validity could be generalized, in other words, if one can apply the findings of a scientific study outside of the study's context (Bryman 2012). In this study, I test a coding scheme inspired by FOI (2020), used for the nature of hate targeting particular people within the journalist-politician-musicians- comedy- and influencer. This partly tests the validity of the coding scheme. The coding scheme and sampling procedure can be outlined clearly by conducting a content analysis, so replications and follow-up studies are attainable (Bryman 2012).

5.4 Sampling and data collection

The collected material consisted in its entirety of 2400 posts, where 600 posts come from each thread. However, comments could contain more than one hateful statement against several persons, and therefore, the number of *statements* will be counted and presented in the material. This is a methodological strength, as hate detection programs can usually only determine a hateful comment, while this study presents all the different types of hostility in the comments. The content of the threads is not static, as the threads are moderated. A thread can have endless posts and be commented on as long as a moderator has not closed it. Some threads only discuss one particular year, while others discuss the reality show ever since it started. The data collection was collected by hand, where all hateful statements were entered into an Excel document and coded according to the coding scheme.

The posts from the four different threads were selected through a simple random sample with an influence of Systematic sample. The threads are found under the head topic “Culture” and sub-topic “celebrity gossip”, where reality shows are discussed. Bryman (2012) describe this as the most basic form of a probability sample. The occasion for being influenced by the human bias is almost zero, and the unit of analysis will not be selected on any subjective criteria because the selection process is random and mechanical. Also, the selection of posts will be made without the author's knowledge. However, I chose to do a Systematic sample of the random samplings with regard to the time frame. This is described by Bryman (2012) as a variation of the simple random sample. Instead of giving all 2400 posts a random number, I only gave the first a random number and included all the following 199 in the material.

Further, to overcome that some of the reality profiles could be targets of hate storms, the samples were also randomized to represent the beginning, middle and end of the thread. For example, different profiles participate at the beginning of the program, and others could be in the show at the end. If just sampling from a specific period of the show, the result would instead tell characteristics of the hate directed to those individuals than the reality profile as a group. Using a random number generator from Random.org, I included the numbers from the Flashback discussion threads, representing posts from 2020.

5.5 Coding

The coding scheme are inspired by FOI (2020), to test the categories in another context. The following categories were the point of departure:

- Origin/ethnicity: Insulting comments about presumed ethnicity, culture, nationality, religion and/or skin colour.
- Sexual harassment: Insult based on the victim's sexual relations or preferences, sexual allusions related to the victim's appearance or behaviour, or an explicit representation of sexual fantasies about the victim.
- Homophobia and Transphobia: Offensive comments about alleged sexual orientation or gender identity.
- Profession/achievement: Insulting comments about career choice or professional achievement.
- Opinions and political ideology: Insulting comments about presumed opinions and/or ideological affiliation.

- Threat: Direct or indirect expression of an intention or desire to harm a person.
- Appearance: Insulting comments about appearance.
- Violations: Other forms of hatred, which do not fall into the categories above. It can touch about everything from slander, nasty rumours, insults and other forms of rudeness.

However, the coding also had an inductive approach, which resulted in the category of profession and religion were not found in the material. The category of hatred due to a person's age and dehumanising content was added after coding the material. Moreover, category number eight was categorised as “Other hateful content instead of violations” because I consider “Violations” to be a too broad term for what is founded in the material. Eight main categories of hateful content were found in the data:

1. Hatred due to a person's age
2. Dehumanizing content
3. Hatred due to a person's known or perceived sexual orientation, summarized into anti-LGTBQIA content
4. Hatred due to a person's physical appearance/look or physical characteristics
5. Sexual abuse, sexual harassments, the spread of sexual rumours or sexual content (e.g., pictures)
6. Directing subtle threats, fantasies about violence against a person, or statements with violent content
7. Xenophobia, dislike, hatred or prejudice against a person's real or perceived origin, skin colour or culture

The posts that could not be coded under any other category, such as gross insults, slander, hateful insults or sexual slurs, were coded under: 8. Other hateful content. The irrelevant information will not be coded. These findings are presented in Table 1.

Table 1. Identified hateful content

Category	Definition	Example of content	Code
Age	Hatred, harassment or insults on the basis of a person's age.		1. Age

Dehumanizing and objectifying content	Content that dehumanizes the person, for example, it resembles animals, extra-terrestrial beings, cartoon characters. Presented as immoral - physically or mentally and become objectified.	“alien”, “pig”, “rat”	2. Dehumanizing
LGTBQIA	Hatred due to a person’s known or perceived sexual orientation, summarized into anti-LGTBQIA content	“nancy-boy”, “faggot”, “dike”, “tranny”	3. LGTBQIA
Physical appearance/look or physical characteristics	Hatred due to a person’s physical appearance/look or physical characteristics	“dismembered”, “disgustingly ugly”, “disgusting snout” “fatso”	4. Physical
Sexual content	Sexual abuse, sexual harassments, spread of sexual rumours or sexual content (e.g. pictures)	“f*ck doll”, “sucking every c*ck”, “h** reward was to f*ck h**”	5. Sexual
Threats	Direct or subtle threats, fantasies about violence against a person or statements with violent content	“shoot them”, “give h** a beating”, “cut h** vocal cord”	6. Threats
Xenophobia	Dislike, hatred or prejudice against a person’s real or perceived origin, skin colour or culture	“n*gger”, “wog”, “monkey”	7. Xenophobia
Other hateful content	Gross insults, slander, hateful insults or sexual slurs	“whore”, “slut”, “f*cker”, “most disgusting participant ever”, “bottom scratch”	8. Other

5.6 Data analysis: Content analysis

To examine the hate and threats in the post and comments in the forum, a research approach suitable for categorizing and analyzing texts' content is sufficient. As the American behavioural scientist Berelson (1952: 18) declare; "*Content analysis is a research technique for the objective, systematic and quantitative description of the manifest content of communication*" (cit. in Bryman 2012: 289). Content analysis is an efficacious technique to analyze documents and texts that allows you to quantify the content in terms of predetermined classes and a systematic and replicable manner (Bryman 2012).

Content analysis can be used for a wide range of materials, such as speeches, books, interviews, and other materials intended to communicate information (Hermann 2008: 151). Objectivity is one of the strengths of content analysis. There is transparency in the procedures for assigning the raw material to categories so that the researchers' preconceptions interfere as little as possible during the process, as rules are specified before dividing their material into different categories. Another quality of this type of method, which also disregards the researcher's personal bias, is that the method is systematic because the rules the researcher uses are consistent. With these qualities, anyone who uses the rules should achieve the same result (Bryman 2012).

Content analysis is concerned with determining the item's apparent content: what it is actually about. Another perspective on content analysis is that it should refer to "specified characteristics". This allows one to investigate what we may call "latent content" or definitions under the surface indicator. Uncovering latent material entails deciphering meanings underneath the surface, such as when a single topic is perceived merely as a source of concern. A similar difference is often made between a focus on the language, such as counting certain words. On the other hand, it focuses on themes within the text, which involves looking for specific ideas within the text (Bryman 2012). This thesis will take the latter approach, as hate, insults and derogatory speech could take various forms of expressions, terms, phrases and language.

5.7 Definition of “hate”

As presented in the previous research and theoretical framework presented, the concept “hate” is a term that most often occurs in hate crimes, online hate, hate sites, hate storms, bullying,

threats, or hate messages. Just as the UN (2019) states, hate is something “*that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are*”. In this study, the term hate is used in the sense of aggressive or abusive communication, but not limited to hatred due to innate group affiliation. However, common "rudeness" will not be included, but the material has only been coded if found in a hateful context, just as Siegel (2020) suggests. For example, calling someone "forty", stupid, or quirky is not enough, but the term must contain a more gross attack against the target.

5.8 Methodological discussion

There are different notions of applying reliability and validity standards to qualitative research (Bryman 2012). There are alternatives to these criteria's; however, I do not find superior alternatives, resulting in an inclusion of reliability and validity. After a thorough analysis of previous studies, researching samples from Flashback's comment boards, and drawing inspiration from FOI's code scheme, I have laid the groundwork for high-quality research to the best possible standard. Moreover, I choose content analysis because of its advantage in studying communication. Additionally, using Flashback as the location, participants or material in the sample is not required to consider the researcher when doing content analysis (Bryman 2012).

One challenge, which can be both an advantage and a disadvantage, for the researcher applying content analysis is that the method is very flexible. There is no unanimous approach to doing content analysis. The particular issue will form the study. However, this also makes the analysis process very interesting (Elo & Kyngnäs 2008). Another argument made by Elo and Kyngnäs (2008) is that one of the most important things to note while reading through the results is to keep the study issue in mind. There were several curious aspects discovered that were unrelated to the purpose of my study. However, where there is an excessive amount of data coded and applied, the data can become unmanageable. Even though analysis should be as impartial as possible, subjectivity must still be considered. Therefore, in this thesis, I reflect on my positionality as a researcher. To research these types of settings, I need to transcend my personal prejudices and gather data in a structured way, which content analysis helps me do.

Managing the data collection entirely manually has both advantages and disadvantages for the study. For example, collecting through a program designed to detect hatred can miss and collect information that is irrelevant to the study. Furthermore, since I have read all the statements, I

have also been able to place the entries in a context and more easily determine if the individual was cynical or not. Furthermore, I did not need to select several people who need to be mentioned in the statements, such as in the FOI (2020) report.

If there were time and space in the study, it would have been interesting to include interviews with people who have been exposed to cyber hatred during and after they have been in a reality show. This would also provide a deeper insight into the hatred expressed on other social media and whether this corresponds to flashback hatred.

A commitment made in the essay is to ensure that several people were mentioned. Almost all participants have been mentioned on at least two occasions so that the hatred is not only directed at one or a few people.

5.9 Ethical considerations

The purpose of this study is not to map the opinions of private individuals but to shed light on the nature of the hatreds, threats and insults to which public individuals are exposed in anonymous online forums. Joinson et al. (2007) argue for when conducting research in online forums, the researcher needs to consider how to protect the users and their anonymity.

Bryman (2012) and Joinson et al. (2007) address the importance of consent between the researcher and the participants. However, both acknowledge that public spaces with large chat rooms and where users are informed that their communication is not confidential or private, the researcher is not obliged to informed consent. Since Flashback is one of Sweden's most visited web pages and open where anyone can read the threads and posts without being a user, I consider the communication on the site is not confidential and, therefore, the ethical concern about consent can, in this case, be neglected. Nevertheless, this thesis sets out to protect the privacy, integrity and anonymity of Flashback users in the best possible way. In a democratic society, individuals, even if they inherit extreme views, has the right to express themselves (within the legal framework). The means to protect citizens' freedoms and rights should not instead become means that restrict them (Tzanou 2010). Therefore following measures have been undertaken in this thesis: none of the usernames is saved in any material, no one has seen the original statements except my supervisor and me, none of the discussion threads will be mentioned, none of the exposed persons will be named, the various reality shows will be

confidential, and no example with a unique character that could be identified via google or the search function will be reproduced (each example is tested in both search engines).

6. Result

The result will be presented as the following:

Firstly, the identified categories will be presented to give an overview of the content of hate. A brief presentation of the categories, with a definition and example of content, will be introduced. Secondly, the distribution of the hate between the categories will be presented, with a summary of the most common features of the content coded in that category. Differences between the four threads will as well be provided. Lastly, the content will be presented in how it is aimed at men and women, resulting in a more in-depth exploration of the hateful material's content.

6.1 The content

In the 2400 comments investigated, about 23% of the comments were hateful. 575 different hateful statements were in total identified in the hostile comments against people who participated in, work with, or connected with the chosen reality shows. Additional 35 hateful statements were directed to people in reality and influencers as a whole. These statements are only included in table 6. Each statement was coded with a maximum of two of the identified categories. Some posts contained hateful content against various persons, and therefore, the number of statements is counted and not the number of hateful comments. A various of people were mentioned, almost all of the participating individuals in the reality shows.

Table 2. The distribution of hateful content in categories

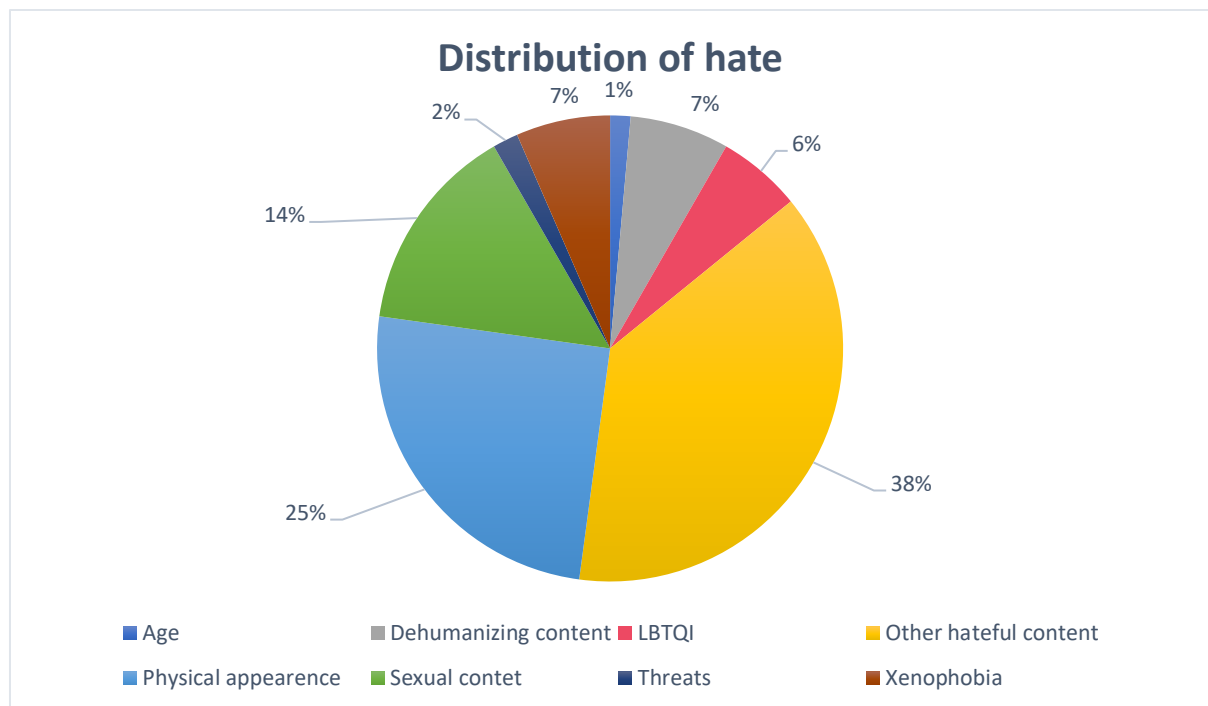


Table 2 shows the distribution of hate in all the hateful posts presented within the identified categories. The smallest amount of categorized content was hatred due to a person's age, most matters of different harassments or derogatory statements about a person's age and/or connection with their participation or lifestyle, which was identified in only 1% of the total content of hate. Threats were identified in 2%. The threats contained a variation of a direct threat of violence, such as killing and injuring people, to the persons' desire to experience violence. In some cases, the author of the post referred to himself/herself when injuring people, while in other cases indicate that "one" or "someone" should harm the person, or imply that the person will be injured or harmed in the future.

6% of the hateful content was identified as hatred due to a person's known or perceived sexual orientation. This contained various anti- LGTBQIA content forms, where slurs and harassment of different sexual orientations were found. The word "faggot" was most commonly used in these statements, closely followed by "nancy-boy".

Xenophobia, along with dehumanizing content, each covers 7% of the total hateful remarks. Xenophobia covered pronouncements that expressed dislike, prejudices or hatred against the targeted person's real or perceived origin, culture or skin colours, such as racist expressions, derogatory terms and stereotypical perceptions. There were expressions of anti-Semitism, Islamophobia along with Afrophobia. However, a more "general" racism aimed for people not

consider as ethnic Swedes. Words as "n*gger" and "wog" repeatedly appeared to depict a person of colour, and Jews were in some cases connected with stereotypical and anti-Semitic notions. Some of the posts connected terrorism to people who in reality are, or believed are Muslims. Dehumanizing expressions often involves the author depriving or denying persons' human attributes and comparing or equating appearance or mentality with animals, cartoon characters, or space creatures.

Further, sexual abuse, sexual harassments, the spread of sexual rumours or sexual content (e.g. pictures) covers about 14% of the material. This category includes a wide range of sexual content, from sexual allusions of people's bodies, fantasies or rumours about people's sex lives to more violent descriptions of sexual acts. Certain parts of the material also include the distribution of nude photos of a person. In various cases, sexual content was coded in addition to other categories. Hatred due to a person's physical appearance/look or physical characteristics occurs in 25% of the material. Various comments concern people's bodies, where words as fat, ugly or disgusting are frequently used to describe a person's physiques. These nouns are also used to describe facial features. Approximately a third of the comments concerning the physical appearance of the statements consider beauty procedures. Further, derogatory terms as "fatso" or "chunk" are frequently applied to a person. "Short" was also used to describe an ugly-looking person, referring to them as "dwarf" in some cases.

A significant majority of the material collected was coded as "other hateful content", which amounted to 38%. This category's content is "general" hate, such as sexism, hateful slander and rumours, derogatory and vicious nicknames, and harmful description of people. Sexual slurs and entitle people as disgusting, whores, sluts and f*ckers, often in a combination, were fairly common, in conjunction with long descriptions of how stupid people are and how low intelligence quotient they possess and compare them to small children's intelligence level. Further, statements, where the author diagnose people with various mental disorders, are included, as well as posts where the author expresses how disgusting people are.

6.2 Men and women

The hateful content differs in several ways, both in qualitative and quantitative manners, when directed to men and women; times of hateful statements, the features of hate, the length of the statements and the author's elaboration of hatred, particularly how graphic descriptions and

comparison are. Table 3 shows the total moments women and men individually received hate in the 2400 comments. Of the total 575 times of hateful statements, women receive about 63,4% of the content, while men receive 36,6%, representing a number of 365 and 210 hateful statements respectively. In this material, women received 155 more hateful statements than men. Thus, hateful content was in 73% more times targeting women than men.

Table 3.

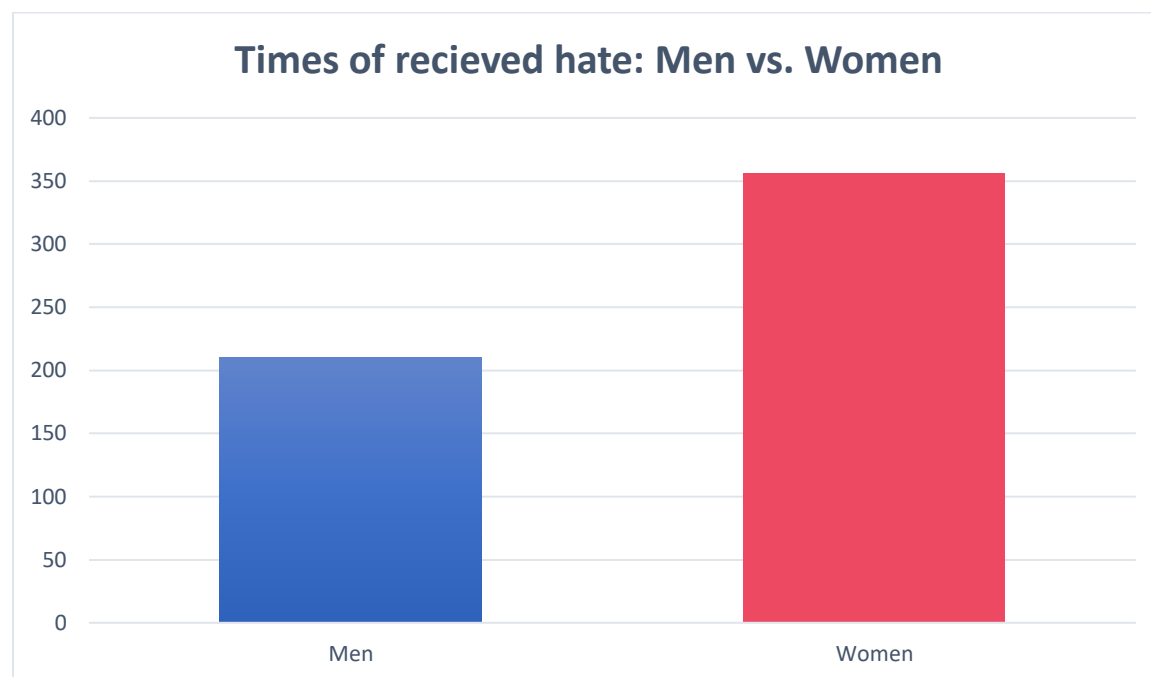
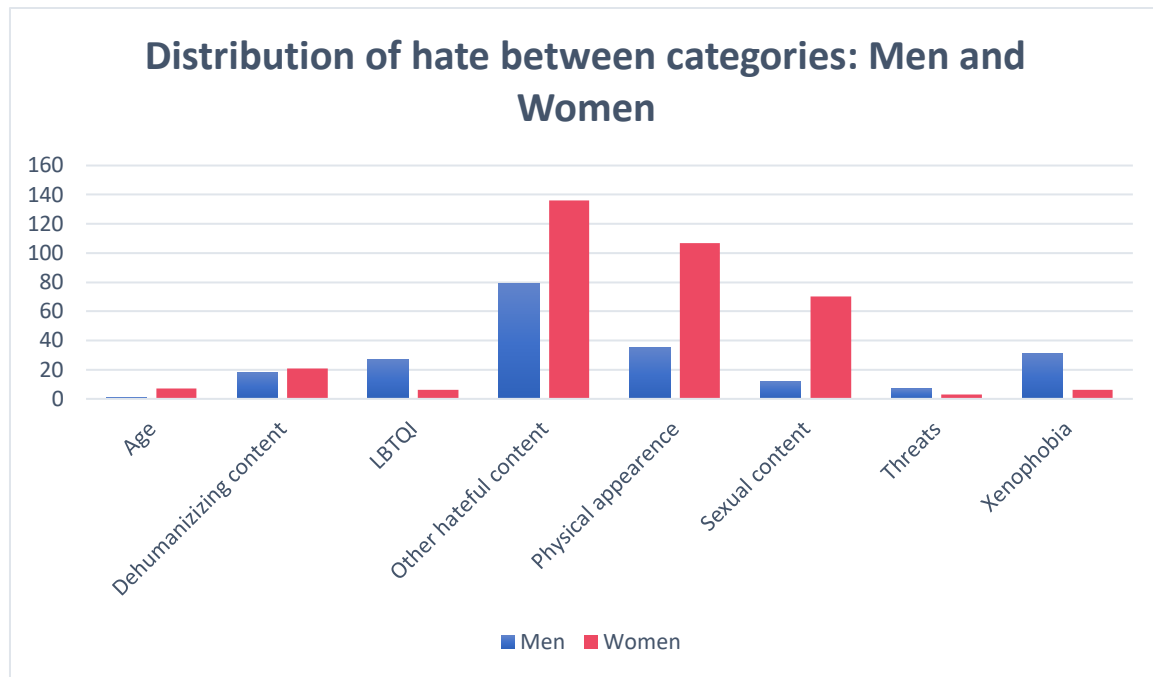


Table 4 shows the nature of the hate for each gender. Women are exposed to hatred due to physical appearance, sexual abuse and sexual harassments, disparaging remarks about a person's age, dehumanising content, and other hateful content; gross offences, slander, hateful insults or sexual slurs. Concurrently, men are to a vaster range targeted by Xenophobia, threats and hatred due to perceived or known sexual orientation, along with almost as much dehumanising content as women. Men were not exposed to hatred based on age, to a lesser degree of sexual abuse or harassments and hatred due to appearance. Other hateful contents are as well a large part of the hatred directed towards men.

Table 4.



6.3 Men and Women – exploring the content

In the category “age”, seven statements against women were found. These statements imply that the women were too old to be in the reality shows, that they should be on a different level in life or have come further in personal development than what one assumes or believes they are, for example, that the woman should be a mother of that age. The statements also contained comments such as that the people were embarrassing, tragic, behaving like children and that the person was disgusting.

The dehumanising content focused both on physical and mental attributes and compared both appearance and mental characteristics with animals, extra-terrestrial beings, cartoon characters and other non-human forms. In several cases, the statements deprive women and men's human qualities, personalities and dignity by presenting or describing them as, for example, immoral. Women are often described with long and more detailed descriptions and are exposed in some posts to comparisons with more than one different non-human beings. As an exemplification, in one post women's appearance, in general, were described and compared to lizards, and one woman, in particular, is described as a lizard and that her face is terrible like a living skull. In several posts, women are equalised with aliens and equated with objects without a human mentality, much like dolls. In one post, the statement's founder literally expresses that a woman is an object that is just "nice" to look at. In other posts, women were described as a duvet cover and a wandering anus. Another characteristic of the dehumanising posts directed at women is

that women are often described as not real women, for the reason that the writer considers that they do not behave like the stereotypical image of how women should act, does not have feminine features or are immoral and do not follow the norm. To clarify this, the authors often use the word women in quotation marks and in some posts, the author compares women in the influencer market and reality show with "real" women.

Simultaneously men within this category are to a greater extent compared with animals and mentally disturbed characters. In two posts, men are equated with cartoon figures whose intelligence and mental capacity are equated with the character. In one case, the word men are used in quotation marks, when asking if there only homosexuals and not "real" men in the show. Men's descriptions only refer to one man, whilst women are in several examples referred to as a group. The descriptions of men are only a sentence at the most, where the illustrations of women most times are more than three sentences long.

The anti- LGTBQIA content contains most of derogatory terms and nicknames for people with a non-heterosexual orientation. The persons with a known or perceived sexual orientation are in all statements interlinked with an offensive term or with a derogatory title, such as "dike", "fag" or "faggot" etc. In several utterances, either women or men were mentioned by name, but instead by terms as "the tail boy", "the dike-cop", "the gay-boy" amongst others. Men received hatred due to sexual orientation about 4,5 times as much as women. However, although men were targets for more hateful content, the nature of hatred did not differ within this category.

Hatred due to a person's physical appearance and physical characteristics contained the longest and most descriptive illustrations of both sexes; nevertheless, women receive more hateful content, longer descriptions of portrayals and statements, more graphic content, and more diverse components of the hate. Women approximately three times as much as men, where hateful statements focused on women 107 times whilst 35 statements target men. A large part of the posts focuses on and describe women's bodies and facial features. Words related to beauty procedures, such as "fillers", "botox", "operation" and "silicone" were found in 31 of the statements targeting women and two times in the content related to men. For example, one post describes a woman who has undergone surgery where the author expresses how the result looks awful and disgusting, comments about her nose, lips and forehead, how plastic and terrible she looks, and how scary and angry she appears.

Some words are well present in the descriptions linked to the operations and the women's appearance. For example, these words are grotesque, disgusting, very ugly, terrible and vulgar together with various swear words, which reinforce the expressions. Other than depict women with different terms like hideous and unattractive, the women's look is also compared with various objects, items or factors. As for instance women's faces are compared with a "traffic accident", "baboon ass", "cheap porn star", "bimbo", "potato", "trash", "jig-saw", "plastic slut", "duck face" and "prostitute". Women's bodies are portrayed with terms as "dwarf", "bottle-shaped", "too small breasts", "too big breasts", "fat", "rag-doll" and "garbage". Additionally, women's buttocks, breasts and lips are the body parts that the opinions focus on the most. Furthermore, some posts focus on women's appearance in general, both women in reality shows as well as women outside the shows. The statements consider women to be more beautiful in the past, and today's women are ugly, operated on, plastic and disgusting. Various posts target all the women in the shows as a group with different terms and descriptions of how ugly they are.

The statements targeting men were often short, containing one or two terms to describe a person's ugliness at most. A limited number of posts focused on facial features, while all the more content focused on the body. Most of the descriptions included terms as "ugly", "fat", "skinny", "childish-looking" and "short". The majority of the post only consisted of that the man in question is unattractive together with various reasons why. The post targeting men only included one man at the time, while the posts targeting women in many cases include more than one. As an illustration of the difference between the nature of the hate follows two examples:

"He's so ugly."

"The girls look really horrible, would not even touch them if I got paid. Bad skin, plastic surgery and sexually transmitted diseases in a wonderful mix."

Sexual content targeted women about 5,8 times as much as men. The sexual harassments and abuses targeting women varied in content: several comments remarked on physical attributed, whilst others were descriptions of presumed sexual activities, sexual desires and fantasies about people's presumed sexual activities. A number of the statements included that the women's sex lives are disgusting, that women are abominable by virtue of their sexual choices, and some of the comments also expressed perceptions that individuals use their body or sexuality to achieve personal benefits. Descriptions of how women like to have sex, graphic illustrations about the women in different sexual contexts, known or perceived statements with whom they have

sexual intercourse with, how many they have sexual intercourse with, and where they have sex are prevalent statements in the material and varies in a scale of violence. For example, a statement could include "she loves to/a good *sexual activity*", "she has slept with half of *place*", "she has slept with over *number of people*", "she is a whore/slut/sex worker. She has sold her body to the whole world".

Additionally, several statements constitute of spreading rumours about people's sex lives and spreading naked or sexualizing images. Few statements discuss that certain of the women work as escorts or are a sex worker. Moreover, the women are called "fuckable" or "non-fuckable" several times, depending on how they look, how they act or the woman's known or perceived lifestyle. Examples of terms used for describing women in the material are: "cock sucker", "whore", "fuck doll", "cheap" and "mattress".

Men are targeted in 12 statements of sexual content. The comments concerning men mainly contained notions about their genitals, that they actually perform non-heterosexual acts in secret and about their assumed or real sex life, sexual rumours. There are no general terms used for describing men. Instead, they are depicted as actors in an illustration of different situations of sexual intercourse. "He likes to/he wants to *sexual act*" and "his *genitals* are *size/proportions/look*" are recurring in the statements. The sexual content targeting men often contain admiration or bear a positive touch, whereas all the statements targeting women have an insulting or degrading character. Only three posts did not describe men as actors. In these statements, the targeted man was described as wanting another man's genitals characteristics or wanting another man to have sexual intercourse with him.

Threats or fantasies about violence against a person targeted man in seven posts and women in three. The threats against men concerned violence against the person in six cases, where desires, fantasies and notions about how the person should be harmed were prominent. Examples of content are; "I hope he gets beaten up every time he leaves his house", "He deserves to be abused" or "I hope someone hit him with a *vehicle*." The last threat was a death threat, was the content illustrates how the author wants to shoot the person. The three threats against women concern death in one case and aggravated assault in the other two cases. One statement expresses how the agent of the statement would strangle one woman. In the other two cases, the author writes that he/she shall force an object into a woman, and in the other case, a wish if someone can cut off the vocal cords of a woman.

Men are targets for xenophobic expressions in almost five times more statements than women, 31 posts and 6 posts, respectively. The N-word occur in seven posts, where six of them are targeting men. For example, one post states that none of the other participants will want to have sex with “that *n-word*” voluntarily. Other racial slurs targeting men are different terms of the word “ape”, “half-breed”, “kebab technician” and “wog”. The other statements described people with different perceived ethnic backgrounds, such as Muslims, Jews and people of colour, with racist stereotypes. Muslims are described as “terrorist”, “wog” and “IS-sympathiser”, where statements claim that the person will commit a terrorist crime and that one do not dare go to airports because of the person. Other posts mean that their behaviour is immoral based on their background, and participants with a background from the Middle East represent behaviour from “the suburbs”.

Other hateful content is the dominant category for both genders. Women received 136 hateful statements that could not be coded under any other category. Men received 79 other hateful statement. 16 of the statements targeting women refer to their intelligence level, ranging from accounts how stupid she is to be comparing their intelligence with young children or animals. Besides those 16, two statements contain the word “retard” in connection to the person's intelligence. Twelve statements consist of the author establishing various mental disorders on people, such as psychopath, borderline or psychopathic traits.

Further, the word “disgusting” is used in fourteen cases to describe a woman, and the words “whore/slut/tramp” are used in eight situations. The latter is often used in scornful contexts, such as “The little tramp is funny to look at”. Also, three statements discussed women's clothing, where they were considered to dress vulgarly - statements that also included the words “disgusting” or “cheap”. Women are also portrayed as bad, disgusting, immoral or even dangerous when associated with feminism. Seven statements concern this, in which the woman in five of the statements is called “femi-nazi”. Four statements discussed if some women's behaviour or look is due to drugs.

Other statements within this category are, for example, that a woman deserves to be mistreated, that it would not be a scandal if a woman had an abortion and “cried out” about it over social media, that one woman should be locked up based on that she is crazy, that one woman has stopped developing and started to phase out, one woman is described as totally useless and on

post discusses how much STDs one woman has, stating that the person "would never dip it in her". One woman is questioned if she has really been involved in alleged sexual abuse, for the reason that; "if you have been sexually exploited, you do not show up on reality TV where you have sex [...]".

Moreover, words found in the discussion about women in the category of "other hateful content" are for example; "cheap", "vulgar", "cunt", "pathetic", "slob" and "unpleasant". Many of the posts contain descriptions of how the author's body reacts to women; "I vomit", "I feel bad", "I twist", "I want to die", etc. are common expressions concerning the discussions the women. To illustrate the content within this category, one representative example follows:

"Is she serious? How disgusting she is!! Really vulgar and cheap. Yuk!!"

In the 79 statements targeting men, 20 statements concern men's intelligence. Similar to the discussion of women, the statements ranged from questioning the man's stupidity, connecting intelligence to behaviour and personality, to comparing the persons intelligence with children or animals. One man's intelligence quotient is discussed, where it is believed that it should be below 80. Additionally, "dumbest participant" or "dumbest person" appeared on several occasions. Expression like "dropped in the floor as new born", different forms of "stupid" and "brain dead" are as well common in the material. As for women, there are statements where the author determine different mental disorder for persons were "Psychopath" or different formulations of "disturbed person" are used in the statements, or that the person in "unpleasant". Other terminology used to discuss men is that they are unsuccessful, tragic, pathetic, pussy, or in some way drug addicted. To illustrated follows an example:

*"*Name* is stupid, backward, tragic, pathetic so disturbing"*

6.4 Hatred due to the group "Reality profiles and influencers".

When the authors of the posts describe influencers and reality profiles as a group, the most common descriptions are about how influencers and reality profiles are "sub-humans", for example by describing them as trash, disgusting, bottom scrapes, disturbed, immoral, unintelligent, mentally handicapped, sad, horrible, brain dead and as inbreeding. One statement expresses that the people within the influencer and reality world is only people from 90a and 00s, "the worst people and the worst bottom scratch". Fourteen comments express derogatory

remarks about the intelligence level of influencers and reality profiles. Further, various comments also contain parables with animals, that influencers and reality profiles behave like monkeys or donkeys. Two comments consider the persons to behave like “wogs”, in a racist derogatory manner. One statement believes that people in the reality and influencer world give birth to “disgusting children”.

Despite statements about influencer and reality profiles looks, as in several statements were depicted as non-f*ckable, horrible looking, trashy looking or just ugly, there were some longer descriptions of the group of influencers and reality profiles in the material. One comment speaks disparagingly of men who adopt more feminine stereotypical actions. The statement says that it is not common for men, for example, to pluck their eyebrows. An excerpt from the comment reads:

*"[...] it's ok to try to be a visual fuckboy forty until 25 [...] if you're still single 30+ and look like a Barbie doll when you put all your money and energy on a fuck*noun* look, that says a lot about your personality. "*

One comment says that the influencer and reality career is really a conspiracy, a coup against the cultural and intellectual level that should be the norm and occupation of the place that society's elite has chosen to withdraw from, so as not to get dirty.

7. Analysis

The aim of the thesis was to outline and categorize the hatred and threats conducted against Swedish reality profiles and influencers on the internet forum Flashback, and examine the possible difference between men and women in the threats and hatred found in the forum. The central research questions for the thesis were firstly, to characterize the online environment and content expressed against influencer- and reality profiles of the chosen reality shows on Flashback forum.

7.1 The nature of hate in the discussion forums

How is hate expressed in the descriptions of reality profiles and influencers in the online discussion threads on Flashback.se?

The material presents a very hostile environment, where hate is expressed in various ways. Just as KhosraviNik and Esposito (2018) describe, the harassments in these case work as a communication and practice in the discussion threads. Flashback offers an interactive and intertextual nature, which works as a channel for people to express their thoughts and ideas of the reality profiles and influencers. The content could be described just as the theory suggests, as a discursive spiral of hate, as almost none of the statements were positive. Both flaming and trolling is apparent to characterizing the essence of the threads, creating hostility. The categories all are very consistent throughout the material, even though they appear in different ways in different threads.

Eight main hateful categories consist of profanity, insults, negative affect, and 'typographic energy' such as capital letters and exclamation marks, all based upon various levels of hostile, aggressive, and offensive or expletive language, targeting or describing an individual. These are primarily found in the categories referring to *what a person is*: their age, dehumanizing content, hatred due to a person's known or perceived sexual orientation, physical characteristics, dehumanizing content or other hateful content involving gross offences, slander, hateful insults or sexual slurs. Further, features of trolling, such as threats of violence, both sexual and physical, are found in the material, categorized mainly as sexual abuse and threats. However, just by examining the coded material, one could not identify that persons try to trick people into a natural or positive discussion to shift the discussion into a more aggressive environment.

Despite that, characteristics of both terms are identified, and together, they create a hostile environment.

Further, in various cases, predominantly in content connected to *how a person looks*, such as xenophobic content, physical appearance and sexual harassment directed to women, as well in the 35 statements directed to the reality profiles and influencers as a whole, people were targeted as a group. As the theoretical framework suggests, when hate becomes a crime, it is commonly directed to individuals because of their group identity and is based upon assumptions of sexuality, gender or religion. Without putting a legal status on the posts, as I believe it requires further theory and knowledge, some statements have similarities of hate crimes, where the violence against one individual works as an act of violence against the whole group. This is illustrated in the material when statements express dislike, prejudices or hatred against Jews with stereotypical and anti-Semitic notions, as well as when women are portrayed as whores and sluts, with different terms and descriptions of women's sex life in general, also, how ugly women both in shows and in general are. Further, there is content in the material with a history of being used to discriminate, insult or dehumanize a group of people, such as monkey, wog, and the n-word. People have used them in history to attack bodies, and in this material, these words are used in a derogatory statement.

All of the tactics for cyber abuse identified by Women's media center are found in the material, where the tactics are corresponding to the main categories found in this material. The spread of sexually explicit images, cyber-exploitation goes under the category of sexual abuse and harassment. Further, gender-based slurs and harassment are coded if an insult as other hateful content, for example, when called slut or whore without context, and if the content is a sexual insult/comment on a woman's body or if /in a context of sexual behaviour, such as f*ckdoll, it was coded sexual abuse or sexual harassment. Sexual objectification is also categorized under the main category of Sexual abuse or sexual harassment, same as slut-shaming. Deadnaming was not per se found in the material; however, the comment about a transsexual person's transformation did somehow become harassment as the author could not be sure that the person in question wants his/her previous identity to be exposed. Flaming was found throughout the material and characterized the language in almost every post. The tactic of cyberbullying, often overlapping with other tactics, are prevalent in the material with its characteristics of insulting, threatening, embarrassing or other types of harassments of people. As shown, there are both textual forms and images.

When the content is directed to physical features, like gender, physical appearance, where a person originates from, the individuals are more occurrent targeted as a group. Simultaneously, it appears that categories of a persons' inside features, as to how a person is or what the individual is, persons are attacked individually. As for the tactics, all of the tactics are often used when harassing women; however, as the findings show, flaming and cyberbullying are the dominant ones when attacking men.

7.2 The nature of hostility when directed at men and women

How does the nature of content directed against women and men differ in the discussion threads at Flashback.se?

The nature of the content differs in several ways when directed at men and women. Except that women may receive more hatred overall, the result shows that statements directed at women are more explicit, derogatory, descriptive, and hostile than when directed at men. As the theoretical framework proposes, by using precise illocutionary and perlocutionary regulation aimed at reducing certain people, legitimizing bigotry against them, advocating violence and hostility, and inducing changes in perceptions and behaviours, hate speech can be used to incite or perpetuate discrimination against women today. It is evident that just as the theoretical framework suggests, a large part of speech directed at women is characterized by misogyny speech. Online abuse is clearly connected to gendered social norms. Women are more targeted by feminine norms, such for their looks and men are more targeted for their behaviours, more corresponding to masculine features. A further example is that a woman should be a mother of a particular age, suggesting that the statement is based on gendered norms. Likewise, in the dehumanizing posts, where women are often described as "not real women" for the reason that they do not behave like the stereotypical image of how women should act (in the author's sense), it also relates to women's vulnerability in the cyberspace based on users' compliance with gendered social norms. This is the case when a statement is asking if there are only homosexuals and not "real" men in the show. The men were considered not to behave like the stereotypical image of masculine men. This could also explain the majority of hate conducted at men in the LGBTQIA-category, as men who do not follow heterosexual norms often are depicted as no real man.

Besides, the theory reveals that misogyny speech consists of profanity, aggressive and sexualized imagery, ad hominem invective, and the framing of coercive sex as an all-purpose correction, all found in the material and the majority of categories and posts targeting women. Women are also called ""fuckable"", and sexual harassments are often coded together with physical appearance actions, and the women's/women's sexual attractiveness is put in rivalry to other characteristics. A woman has to be beautiful, act in a way the author thinks is good or ok, and should not have a known sex life to be ""fuckable"". Women who do not fulfil these criteria/criteria's are perceived to be a whore, f*ckdoll or mattress. Men, on the other hands, are described as ""better"" individuals in sexual content and often as actors, while women are recipients.

Just as men and women are populating the cybersphere in different ways, where women should be known to inhabiting the cybersphere as a precarious subject, women and men are described and targeted in the same ways in the discussion threads. Therefore, when harming men and women can result in different hateful expressions. Further, a woman who participates in a reality show or post thing on Instagram that could be considered a violation of gender ideology's social norms, people are triggered and target the person with harmful, often sexualized speech and objectified comparisons her in place. As the findings show, women are in several statements compared with various objects, items or factors, in one case even referred to as an object. As in line with previous studies, women are more often targeted by sexual harassment and insults of their physical appearance, with other content often referring to physical attributes. Men more often by assaults linked with their psychological attributes. These are specially outlined within the dehumanizing and physical appearance categories. In the dehumanizing category, men are to a greater extent compared with animals and mentally disturbed characters, and women with funny-looking characters. In the physical appearance category, the description of women was long, gross, compared with multiple objects and portrayed as with different shapes; meanwhile descriptions of men contained one or two terms to describe unattractiveness.

It is apparent that the content consists of various gender-based harassments. In advance, the material shows that gender-based violence online is prevalent and takes both sexual and psychological forms. The categories more common for women have closely linked to the threats of gender-based violence women experience in society and real life, like unwanted, offensive or sexually explicit statements, sharing intimate images or videos, incitement of violence, stalking, harassments. Just as the theoretical framework suggests, much of the content directed

at women could be recognized as misogyny speech. The content categorized under both sexual abuse and physical appearance draws on perceptions of how women acted and looked back in time, and when not living up to the perceived standards, they were described as immoral. For instance, in the sexual abuse and sexual harassment category, only two comments were derogatory against men, whilst all of the comments directed to women were insulting, exploiting, or derogatory. This shows that the sexual content directed at women takes a sexist form, and women's bodies and sexuality are used to degrade them, while the sexual harassment targeting men in some ways are used to raise them. This could be explained just as theories imply that this kind of speech is a form of violence against females, a form of violence that reinforces and aggravates discrimination against women. Further, these notions might be based on the principles of the persistent unequal power relations between women and men and the long-term effects of structural gender inequity. Targeting men and women in this way might possibly be based on the principles of the persistent unequal power relations between women and men, and the way of targeting the genders in different ways could work as a form of violence that reinforces and aggravates discrimination against women.

8. Discussion and conclusion

Looking back at the previous research and the presented theoretical framework, a significant part agrees relatively well and some different and new results. There are significant similarities with the FOI (2020) study, where the findings show that women are exposed to more hatred than men and the fact that hatred of women and men looks different; women may receive more sexual harassment and appearance-related insults while men are more often insulted due to mental characteristics. Further, the coding scheme seems to be a good basis for conducting a content analysis in anonymous discussion environments online. From a theoretical point of view, this study has contributed to categories that could be common against public figures in these forums and what people who intend to participate in a reality show or become an influencer can expect to depict them. Sexual harassments are a problem for people identified as reality profiles or influencers, together with racism and homophobia, and future research needs to focus on this issue.

This thesis show that when people discuss reality profiles and influencers in a forum, the language, expression and characteristics tend to be incredibly hostile. Additionally, there is a gendered form of hate expressed in these threads. Despite increasing equality between the genders, women are still victims of hatred and violence in a different way than men. Both this thesis and previous research imply that when women take a more notable place in society, appear in places they have not previously seen in, or violate stereotypical norms, people cruelly depict them (see Barak 2005; Bartow 2009; Citron 2009; Heil and Piskorski 2009; Ritter 2009; Franks 2012; Gardiner 2018; ConnectAmericas 2021). Equally, as in line with Kang (2000), Awan (2014) Awan (2016), and Pantti et al. (2019), minorities who do not have stereotypical Western skin colour or religion are subject to racist statements, prejudices and slurs, such as calling them "ape", "half-breed", "kebab technician" and "wog", and describe known or perceived Muslims as "terrorist" and "IS-sympathiser". Derogatory expressions as "dike", "fag" or "faggot" in the material shown as well that people with a non-heterosexual identity are targets for anti- LGBTQIA content, even though those rights have also received more attention. This insight is very interesting, as it partly tells about a hidden side of the users' attitude to equality. One may ask how these attitudes and expressions about women, people of color and non-heterosexuals can be so present online in a democratic country like Sweden. This also tells something about how freedom of speech in a democratic country like Sweden can be used against other people.

Framing this issue within the notion of the disinhibition effect theory, just like the SIDE model, can imply why and which conditions for users to express themselves in a certain way. However, this could not be stated in this thesis, but this is an interesting notion for future research in this area. As there are 610 identified various hostile statements found in the material, the environment, in line with Suler (2004), could be characterized by toxic disinhibition and Social Identity Model of Deindividuation Effects (Lea et al. 2001). The amount of hate, the way of expressing and the very gross insults found in the material suggest that these two effects could affect people's expression in specific ways. Just as these two theories suggest, the material consists of hate, threats, harsh criticism, slander, the spread of intimate pictures, abuse and harmful content. One could ask if persons would express themselves in a public discourse, where these environments do not exist. In creating the specific nature in this environment, all of the six factors of the disinhibition effect could contribute to the hostile settings and the deindividualization of oneself and others. Flashback is anonymous, and the users use nicknames and pictures of all, from celebrities to animals. Therefore, calling people sluts, idiots, f*ckdolls together with other threats and harmful posts does not affect the real-life person. This reflects the strategic side of SIDE as well, where anonymity enables individuals to express opinions that non-anonymous and public contexts would have to be more lenient.

Due to transmitted threats or other degenerate acts found in the material, the authors may feel they can absolve themselves of responsibility for such behaviours. The person they are insulting, threatening, or sexual harassing does not know who they are. As the threats also consisted of death-threats, calling people racial slurs or dehumanize a person, a territory that people not usually visit in the outside world or the public discourse, the authors of these posts are shields from others' body language, facial expressions or other expressions when uttering these negative statements. When people refer to others, to the real-life persons when indicating that "one" or "someone" should harm the person or imply that the threats will be accurate in the future, it could also result from these two effects. The "what happens online has nothing to do with the reality" could explain this very hostile environment and nature inside the discussion forum.

Moreover, where anonymity reigns supreme, the forum member's identity takes precedence over personal identity, as to the SIDE model's cognitive side. Following these rules is more prominent if the Flashbacks norms and other users' behaviour are harmful, as individuals

associate more deeply with the community than their real-world identities. Calling people whore, fatso, alien, or bottom scratch is not common in an environment not characterized by these features. The asynchronicity also comes into place, as statements on Flashback are posted without immediate response or a response at all. After posting a hateful post, there is no need to return to the thread to read replies. People in this environment also write things that in real-world, where anonymity or invisibility could not hide them, would be criticized, punished or in some way retribution would be of question. It seems that the respect for status and authority is minimal and almost as a competition of who authors the most comprehensive and most baggy statements. This is also shown when people are attacking several persons in one statement. However, as previously stated, this is not something that can be established or perceived as conclusions. Nevertheless, it tells something about the problem and is an exciting area to research further. Studies that would be very interesting and may answer these questions are interviews or surveys aimed at people who express their hatred in these discussion threads.

One could ask if the online environment and discussion forums such as Flashback have been the new arena to express this type of hatred. The threads are influenced by both disinhibition and deindividualization, and there is no chance to identify the person behind the posts. An example of this is the use of the n-word, which one does not hear much of today in the public discourse, other than online. As reality profiles and influencers were not studied in depth before, this study's findings show that influencers and reality profiles seem to be hated in a similar way as previously researched groups, such as journalists, musicians, politicians, etc. However, this result presents more sexual abuse and harassments against both genders, but mainly against women. This thesis is somehow unique in the way it shows de facto expressions conducted at individuals. Whether one has a notion of what is commonly expressed on the internet or no notion at all, I believe that this study's result is disturbing. As the theory suggests, this is a form of violence and discrimination of mainly women and men in a new arena. In this material, threats were not as frequent as has been shown in previous research. This could depend on the data collection method, as threats can be something one would instead write in person to the targeted victim, then in a thread where the person cannot see it. It also shows what type of hate speech that targets victims in an environment influenced by toxic disinhibition and deindividualization. One finding not in line with Nadim and Fladmoe (2019) is that women receive such a majority amount of hate. However, many studies do not explore the differences between genders in that particular way. Men receive a more significant deal of xenophobia,

threats, and hatred due to perceived or known sexual orientation not identified by any of the included previous research in this thesis, other than FOI:s memo from December 2020.

For future research, other social discussion forums would be intriguing to study. Further, it would be interesting to test the coding scheme developed in this study on other discussion threads or in other forums to see if the content in, for example, commentary fields are similar to Flashback. A comparison between different forums could be interesting to see if the disinhibition effect and SIDE are as prevalent for the users, or if not, the content will differ. Furthermore, this study does not capture if, as Gardiner (2018) findings show, if, for example, women of a minority are a target in a broader range or categorized by multiple hateful categories. To study this, one needs to map all the people mentioned. As this thesis shows, influencer and reality profile is also a group of very young people who do not have any forms of protection that, for example, people in a public sector have, a group exposed to very crude, sexual and degrading content. Therefore, this population needs to be the subject of further study, not just social research.

References

Books

- Ahmed, S. (2014). *"The cultural politics of emotion"*. (Second edition.) Edinburgh: Edinburgh University Press.
- Bryman, A. (2012) *"Social Research Methods"*. 4th ed. Oxford: Oxford University Press.
- Darian-Smith, E and McCarty P.C. (2017). *"The Global Turn: Theories, Research Designs, and Methods for Global Studies"*. (1st ed.).
- Fredman, S. (2018). *"Comparative human rights law"*. Oxford University Press.
- Hermann, M. (2008). *"Content Analysis"* in Klotz, A., and Prakash, D. (eds.) *"Qualitative Methods in International Relations: A Pluralist Guide"*. Basingstoke: Palgrave MacMillan.
- Hine, C. (2015). *"Ethnography for the Internet: embedded, embodied and every day"*. London: Bloomsbury Academic.
- Ritzer, G. (2011). *Globalization: the essentials*. Chichester: Wiley-Blackwell.

Chapters

- Hermann M.G. (2008) *"Content Analysis"*. In: Klotz A., Prakash D. (eds.) *"Qualitative Methods in International Relations"*. Research Methods Series. Palgrave Macmillan, London. https://doi.org/10.1057/9780230584129_10
- Jane, Emma A. (2020). *"Online abuse and harassment"*, in Karen Ross (eds.). *The Encyclopedia of Gender, Media and Communication*. Wiley-Blackwell
- Kopytowska, M. (2017). *"Introduction: Discourses of Hate and Radicalism in Action"*. In Kopytowska, M. (ed.), *Contemporary Discourses of Hate and Radicalism across Space and Genres* 1–12. Amsterdam: John Benjamins.
- Parekh, B. (2012). *"Is There a Case for Banning Hate Speech?"* In M. Herz and P. Molnar (Eds.), *"The Content and Context of Hate Speech: Rethinking Regulation and Responses"* (pp. 37-56). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139042871.006

Ritter, B.A. (2009). "*The new face of sexual harassment*". In Torres Coronas, T and Arias Oliva, M., (Eds.), *Encyclopedia of human resources information systems: Challenges in e-HRM* (pp. 655–660). Pennsylvania: IGI Global

Royzman, E. B., McCauley, C., & Rozin, P. (2005). "*From Plato to Putnam: Four ways to think about hate*". In R. J. Sternberg (Ed.), *The psychology of hate* (pp. 3–35). Washington, DC: American Psychological Association.

Tynes, B. (2005). "*Children, adolescents and the culture of online hate*," in *Handbook of Children, Culture and Violence*, (eds.) N. E. Dowd, D. G. Singer, and R. F. Wilson (Thousand Oaks, CA: Sage), 267–289. doi: 10.4135/9781412976060.n14

Articles

Aumer-Ryan, K., and Hatfield, E. (2007). "*The design of everyday hate: A qualitative and quantitative analysis*". *Interpersona*, 1(2), 143–172.

Awan, I. (2014). "*Islamophobia and Twitter: A Typology of Online Hate Against Muslims on Social Media*". *Policy & Internet*, 6(2), 133-150.

Awan, I. (2016). "*Islamophobia on Social Media: A Qualitative Analysis of the Facebook's Walls of Hate*". *International Journal of Cyber Criminology*, 10(1), 1-20.

Baider, F. (2018). "*Go to hell fucking faggots, may you die!*" *framing the LGBT subject in online comments*. *Lodz Papers in Pragmatics*, 14(1), 69-92. <https://doi-org.ezproxy.ub.gu.se/10.1515/lpp-2018-0004>

Barak, A. (2005). "*Sexual harassment on the internet*". *Social Science Computer Review*, 23(1), 77–92.

Bartow, A. (2009). "*Internet defamation as profit centre: the monetization of online harassment*". *Harvard Journal of Law and Gender*, 32(2), 383–429.

Blaya, C. (2019). "*Cyberhaine*", *Les Jeunes et la Violence sur Internet*. Paris: Nouveaux Mondes.

Bolognesi, M., Pilgram, R., & van den Heerik, R. (2017). "Reliability in content analysis: The case of semantic feature norms classification". *Behavior Research Methods*, 49(6), 1984–2001. <https://doi.org/10.3758/s13428-016-0838-6>

Chrisopherson, K. M. (2007). "The positive and negative implications of anonymity in Internet social interactions: "On the Internet, Nobody Knows You're a Dog". *Computers in Human Behavior*. 23:(6) 3038–3056.

Citron, Danielle Keats (2009b). "Law's expressive value in combating cyber gender harassment". *Michigan Law Review*, 108(3), 373–415.

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, E.H and Quattrociocchi. (2016). "The spreading of misinformation online". *Proceedings of the National Academy of Sciences*, 113(3), 554-559.

Elo, S, Kyngnäs, H (2008). "The qualitative content analysis process". *Journal of Advanced Nursing*62(1): 107-115

Fitness, J., and Fletcher, G. J. O. (1993). "Love, hate, anger and jealousy in close relationships: A prototype and cognitive appraisal analysis". *Journal of Personality and Social Psychology*, 65(5), 942–958.

Foxman, A. H., and Wolf, C. (2013). *Viral Hate: Containing its Spread on the Internet*. New York, NY: Macmillan.

Franks, Mary Anne (2012). "Sexual harassment 2.0". *Maryland Law Review*, 71(3), 655–704.

Fridh, M. (2018). *Bullying, violence and mental distress among young people. Cross-sectional population-based studies in Scania, Sweden*. Lund University: Faculty of Medicine.

Frijda, N. H. (1986). "The emotions: Studies in emotions and social interaction". Cambridge, UK: Cambridge University Press.

Gardiner, B. (2018). "It's a terrible way to go to work:" what 70 million readers' comments on the Guardian revealed about hostility to women and minorities online. *Feminist Media Studies*, 18(4), 592-608.

Halperin, E. (2008). "*Group-based hatred in intractable conflict in Israel.*" *Journal of Conflict Resolution*, 52(5), 713–736.

Hara, N and Huang, B. (2011). "*Online social movements*". *Annual Review of Information Science and Technology*, 45(1), 489-522.

Heil, B, and Piskorski, M. (2009). "*New Twitter research: Men follow men and nobody tweets*". *Harvard Business Review*. Available at:
http://blogs.hbr.org/cs/2009/06/new_twitter_research_men_follo.html).

Hinduja, S., & Patchin, J.W. (2010). "*Bullying, cyberbullying and suicide*". *Archives of Suicide Research*, 14(3), 206-221. doi:10.1080/13811118.2010.494133

Jane, E. (2015). "*Flaming? What flaming? The pitfalls and potentials of researching online hostility*". *Ethics and Information Technology* 17(1). 65–87

Joinson, A and Pain Schofield, C. (2007). "*Self-disclosure, Privacy and the Internet. The Oxford handbook of Internet psychology*". 2374252. 10.1093/oxfordhb/9780199561803.013.0016.

Kahn, R and Kellner, D. (2004). "*New Media and Internet Activism: From the 'Battle of Seattle' to Blogging*". *New Media & Society*, 6(1), 87-95.

Kansara, K.B., and Shekokar, N. (2015). "*A Framework for Cyberbullying Detection in Social Network*". Vol.5, No1.

Kang, J. (2000). "*Cyber-race*". *Harvard Law Review*, 113, 1130-1208.

Kaul, V. (2012). *The Changing World of Media & Communication*. *J Mass Commun Journalism* 2:116. doi:10.4172/2165-7912.1000116

Keipi, T., Näsi, M., Oksanen, A., and Räsänen, P. (2017). *Online Hate and Harmful Content: Cross-National Perspectives*. Abingdon and New York: Routledge.

KhosraviNik, M., and Esposito, E. (2018). “*Online hate, digital discourse and critique: Exploring digitally-mediated discursive practices of gender-based hostility*”. *Lodz Papers in Pragmatics*, 14(1), 45-68.

Lea, Martin, Russel, Spears, de Groot, Daphne. (2001). “*Knowing me, knowing you: anonymity effects on social identity processes within groups*”. *Personality and Social Psychology, Inc.* Vol. 27:(5) 526–537.

Leung, A. N., Wong, N., and Farver, J. M. (2018). “*You are what you read: the belief systems of cyber-bystanders on social networking sites*”. *Front. Psychol.* 9:365. doi: 10.3389/fpsyg.2018.00365

Maireder, A and Schwarzenegger, C. (2012). “*A MOVEMENT OF CONNECTED INDIVIDUALS: Social media in the Austrian student protests 2009*”. *Information, Communication & Society*, 15(2), 171-195

Mattoni, A., and Treré, E. (2014). “*Media Practices, Mediation Processes, and Mediatization in the Study of Social Movements*”. *Communication Theory*, 24(3), 252-271.

Nadim, M., and Fladmoe, A. (2019). *Silencing Women? Gender and Online Harassment*. *Social Science Computer Review*. <https://doi.org/10.1177/0894439319865518>

Oatley, K., and Jenkins, J. M. (1996). “*Understanding emotions*”. Malden, MA: Blackwell.

Pantti, M., Nelimarkka, M., Nikunen, K., and Titley, G. (2019). “*The meanings of racism: Public discourses about racism in Finnish news media and online discussion forums*”. *European Journal of Communication*, 34(5), 503-519.

Pelzer, B., Kaati, L, and Akrami, N. (2018). “*Directed Digital Hate*” *IEEE International Conference on Intelligence and Security Informatics (ISI)*, Miami, FL, 2018, pp. 205-210, doi: 10.1109/ISI.2018.8587396.

Perry, B., & Olsson, P. (2009). “*Cyberhate: The globalization of hate*”. *Information & Communications Technology Law*, 18(2), 185-199.

Selkie, E., Kota, R., Chan, Y., and Moreno, M. (2015). “*Cyberbullying, Depression, and Problem Alcohol Use in Female College Students: A Multisite Study*”. *Cyberpsychology, Behavior, and Social Networking*, 18(2), 79-86.

Shaw, L. (2012). *“Hate Speech in Cyberspace: Bitterness without Boundaries”*, 25 Notre Dame J.L. Ethics & Pub.

Staude-Müller, F., Hansen, B., & Voss, M. (2012). *“How stressful is online victimization? Effects of victim's personality and properties of the incident”*. European Journal of Developmental Psychology, 9(2), 260-274.

Sternberg, R. J. (2003). *“A duplex theory of hate: Development and application to terrorism, massacres, and genocide”*. Review of General Psychology, 7(3), 299–328.

Suler, J. (2004). *“The Online Disinhibition Effect”*. CyberPsychology & Behavior, 7(3), 321-326.

Tzanou, M. (2010). *“The EU as an emerging 'Surveillance Society': The function creep case study and challenges to privacy and data protection”*. Vienna Journal on International Constitutional Law, 4(3), 407-427. <http://www.icl-journal.com/>

Reports

Barton A, and Storm, H. (2014). *“Violence and harassment against women in the news media: A global picture,” International women's Media Foundation and International News Safety Institute*. <https://16dayscampaign.org/wp-content/uploads/2019/03/Violence-and-Harassment-Against-Women-in-the-News-Media-PDF.pdf>

Bladini, M. (2017). *”Hat och hot på nätet : en kartläggning av den rättsliga regleringen i Norden från ett jämställdhetsperspektiv”*. Retrieved from <http://urn.kb.se/resolve?urn=urn:nbn:se:norden:org:diva-4896>

Council of Europe (2017). *“Gender Equality Strategy”*. Implementation of the Council of Europe Gender Equality Strategy 2014-2017 Annual report 2017. <https://rm.coe.int/annual-report-2017-gender-equality-strategy-2014-2017/168077c5e0>

Eggebo, H. and Stubberud, E., (2016). *”Hatfulle ytringer, Delrapport 2: Forskning på hat och diskriminering”*. Oslo: Institutt for samfunnsforskning

Eriksson Almgren, T., Ekblad, M. and Sjören, T. (2017). "Hot mot kvinnliga opinionsbildare". Kantar SIFO. <http://tu.se/wp-content/uploads/2017/04/PDF-Hot-mot-kvinnliga-opinionsbildare.pdf>

European Union (2014): "Violence against women: an EU-wide survey" European Agency for Fundamental Rights (FRA). ISBN 978-92-9239-342-7, available at: <https://www.refworld.org/docid/5316ef6a4.html>

Fernquist, J., Kaati, L., Asplund Cohen, K., Pollack Sarnecki, H., Pelzer, B., Akrami, N and Lindberg, S. (2020). "Det digitala hatets karaktär. En studie av hat mot kvinnor och män i utsatta yrkesgrupper". Totalförsvarets forskningsinstitut (FOI).

Guterres, A. (2019). "United Nations Strategy and Plan of Action on Hate Speech." https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_EN.pdf

Journalistförbundet. (2014). Enkät om kränkningar och påtryckningar mot journalister. Stockholm: Journalistförbundet.

Kaati, L., Asplund Cohen, K., Akrami, N., Shrestha, A., Fernquist, J., Pelzer, B., Welanders, F., & Isbister, T., (2019). "Digitala Slagfält. Totalförsvarets forskningsinstitut". FOI-R--4813--SE

Organization for Security and Cooperation in Europe (OSCE) / Office for Democratic Institutions and Human Rights (ODIHR). (2009). "Hate crime laws: A practical guide". <http://www.osce.org/odihr/36426>

Plan International (2020). "Free to be online? Girls' and young women's experiences of online harassments". The State of the World's Girls 2020. <https://plan-international.org/publications/freetobeonline>

News paper

Amiri, F. in NBC News (2019). "The World Wide Web is 30 years old — and its inventor has a warning for us". March 12, 2019, 5:17 PM CET

Murakami, S and Park, J. in Reuters (2020). “*Death of reality TV show star in Japan spotlights cyber bullying*”. <https://www.reuters.com/article/uk-japan-internet-terracehouse/death-of-reality-tv-show-star-in-japan-spotlights-cyber-bullying-idUKKBN23A03O>

Winterburn, T. in Euro Weekly News (2021). “*Fellow WAG Hints Cyber-Bullying Drove Friend Kasia Lenhardt To Suicide After Splitting From Jerome Boateng*”. <https://www.euroweeklynews.com/2021/02/12/fellow-wag-hints-cyber-bullying-drove-friend-kasia-lenhardt-to-suicide-after-splitting-from-jerome-boateng/>

Internet sites

ConnectAmericas (2021). “*Women are driving the social media revolution*”. <https://connectamericas.com/content/women-are-driving-social-media-revolution> (2021-02-02)

European Commission (2021). “*Gender-based violence (GBV) by definition*”. https://ec.europa.eu/info/policies/justice-and-fundamental-rights/gender-equality/gender-based-violence/what-gender-based-violence_en (2021-02-14)

Flashback (2021). <https://www.flashback.se> (2020-12-01)
<https://www.flashback.se/flashbacks-historia> (2020-12-12)

Influencer Marketing Hub (2021). <https://influencermarketinghub.com/what-is-an-influencer/>

NodeGraph (2020). “*How much data is on the internet? The Big Data Facts Update 2020*”. Mars 26th. <https://www.nodegraph.se/how-much-data-is-on-the-internet/> (2020-11-28)

Women’s media center (2021). “*Online Abuse 101*”. VMC Speech Project. <https://www.womensmediacenter.com/speech-project/online-abuse-101/> (2021-01-18)