NORWEGIAN
INSTITUTE FOR
SOCIAL RESEARCH

Report 2016:20

Hate speech, report 1

Research on the nature and extent of hate speech

Marjan Nadim and Audun Fladmoe

Norwegian Institute for Social Research
Report 2016:20

Hate speech, report 1

Research on the nature and extent of hate speech

Marjan Nadim and Audun Fladmoe

Translated to English by Samtext Norway AS

Report 2016:20

Norwegian Institute for Social Research
Munthes gate 31
PO Box 3233 Elisenberg
NO-0208 Oslo, Norway

ISBN (online): 978-82-7763-539-2
ISSN (online): 1891-4314

www.samfunnsforskning.no

Content

Preface

The Institute for Social Research (ISF) has been commissioned by the Norwegian Directorate for Children, Youth and Family Affairs (Bufdir) to prepare three reports on hate speech. ISF has cooperated with KUN[1] and lawyer Jon Wessel-Aas in the preparation of these reports. The background for the project is the strategy to prevent hate speech presented by the Norwegian Government in November 2016. The reports are included as part of the knowledge base for this strategy.

Marjan Nadim from ISF has been Project Manager. The members of the project team have worked together in one group, but divided the work on the reports among themselves. For Report 1, Marjan Nadim and Audun Fladmoe, also from ISF, have reviewed relevant statistics and research on the nature and extent of online hate speech. For Report 2, Helga Eggebø and Elisabeth Stubberud (KUN) have reviewed research that sheds light on the relationship between hate speech and discrimination, bullying and violence. For Report 3, Jon Wessel-Aas has investigated the legal boundary between freedom of speech and protection against hate speech, while Audun Fladmoe and Marjan Nadim have described ongoing discussions about where such boundaries should be drawn.

Simultaneously with this project, ISF and Jon Wessel-Aas have also been working on a project for the Ministry of Justice and Public Security relating to the prevention of online hate speech and hate crime. The projects have several common factors, particularly in relation to parts of the literature review and legal investigations. This has allowed the project group to gain in-depth knowledge of the research literature, but also implies that there is somewhat of an overlap between the reports prepared for the Norwegian Directorate for Children, Youth and Family Affairs and for the Ministry of Justice and Public Security respectively.

1 KUN is a private foundation located in Steigen in Nordland County, Norway that works with gender equality, diversity and integration. For more detailed information, go to www.kun.no.

Our sincere thanks to Joseph Vasquez, Karen Sofie Pettersen and Cecilie Håkonsen Sandness at the Norwegian Directorate for Children, Youth and Family Affairs and Kari Steen-Johnsen, Arnfinn H. Midtbøen, Marte Winsvold and Bernard Enjolras at the Institute for Social Research for their helpful input on previous drafts of the reports. Jon Haakon Hustad at the library at ISF has provided invaluable help with the literature search.

Oslo and Steigen, 30 September

Marjan Nadim (Project Manager)
Audun Fladmoe
Helga Eggebø
Elisabeth Stubberud
Jon Wessel-Aas

Summary

Authors Marjan Nadim and Audun Fladmoe

Title Hate speech, report 1: Research on the nature and extent of hate speech

Summary The purpose of this report is to gather research-based knowledge concerning:

- the extent of online hate speech
- which groups in society are particularly subjected to online hate speech
- who produces hate speech, and what motivates them

Hate speech is commonly understood as any speech that is persecutory, degrading or discriminatory on grounds of the recipient's *minority group* identity. To be defined as hate speech, the speech must be conveyed publicly or in the presence of others and be directed at a certain group or an individual's (assumed) group identity. The concept pertains notably to particularly vulnerable minority groups. Section 185 of the Norwegian Penal Code provides legal protection against discrimination or hate speech on grounds of: a) skin colour or national or ethnic origin, b) religion or life stance, c) homosexual orientation or d) disability. The concept of "hate speech" is, however, often used in a broader sense than what ensues from the Norwegian Penal Code. The present report discusses research that concerns both criminal and non-criminal hate speech, as well as other offensive statements or displays that might be *perceived* as hate speech.

The report demonstrates that statistics on officially reported instances of alleged criminal hate speech are extremely limited and not adequate for determining the extent of hate speech. Various questionnaire-based surveys serve to fill out the picture, but one of the problems of this type of research is that very few studies have investigated the extent of hate speech directly. Different studies have applied different definitions of hate speech and other offensive statements and hence arrived at greatly differing figures for the extent. Many of the studies have primarily addressed online hate more generally, without reference to the concept of "hate speech". This means that there are few studies that examine hate speech aimed at specific group identities.

Regardless of how the prevailing terms have been defined in the empirical studies, our systematic review shows that ethnicity and sexual orientation are the commonest grounds for hate speech. Less research has been conducted on online hate targeting disability. Furthermore, young people are far more often subjected to offensive experiences online than others. Overall, gender differences are relatively insignificant when it comes to exposure, but women and men are subjected to different forms of online hate.

Analyses of a Norwegian survey conducted in June 2016 show that around two per cent of the population respond that they have been targets of hate speech in social media on the grounds protected by Section 185 of the Norwegian Penal Code. Disregarding which grounds that hate speech is actually targeting, our analyses show that around seven per cent of the Norwegian population reported being the targets of what they themselves *perceive* as hate speech. However, most of these instances of hate speech target grounds that are not covered by Section 185.

The report also cites analyses of surveys conducted in 2013 and 2015 among persons of immigrant descent from Eastern Europe, Asia and Africa, and among journalists. These surveys show that immigrants are more exposed than the majority population to offensive and derogatory comments based on the grounds provided for by Section 185, but that the majority population has more frequently been subjected to offensive comments on other grounds. Journalists are generally far more exposed than the general population to offensive or derogatory comments and threats. These comments are, however, typically based on grounds that are *not* protected by Section 185.

Moreover, the literature review reveals that the perpetrators of online hate speech are typically men, and that men have a greater tolerance of internet hate than women do. The perpetrators are also often motivated by factors other than a strong feeling of hate. Instead, factors such as thrill-seeking and an internet culture of defamatory language are more likely the key to understanding the underlying dynamics of hate speech. That said, the victims of hate speech are not random. The perpetrators base their views on prejudices, stereotypes and subjective assumptions about differences between groups in society, and hate speech is directed at those perceived as “different”.

Based on the literature review in this report and Report 2 (Eggebø & Stubberud 2016), we identified four research needs:
1) research on the extent and experience of hate speech, including discriminatory and offensive statements and displays;
2) research on the perpetrators of hate speech;
3) textual analysis of hate speech and discriminatory statements in the public sphere and
4) research on the consequences of hate speech.

Index terms Hate speech, Internet, online bullying

1 Introduction

Hate speech topicalises the tension that exists between, on the one hand, freedom of speech and the protection of minorities on the other. These are two central values in a liberal democracy, but the boundaries between them are not a given. Different countries have not surprisingly balanced these values differently over time, but the general trend has been for hate speech to be brought under increasingly more stringent regulation (Bleich 2011). While anti-Semitic hate speech was not uncommon in the media in the interwar period, in most European countries this is now prohibited by law.

In parallel with a reduction in hate speech in conventional media, the emergence of the internet and social media has shifted public tolerance limits. Not least with the popularity of social media, the scope for making comments that go unmoderated in the public sphere has greatly increased. The increasing use and accessibility of social media, combined with their networked structure now mean that hate speech can be spread rapidly and reach more people. In this way, the scope for targeting individuals and groups with hate speech has grown enormously.

There is broad consensus that hate speech is a significant problem for society. Hate speech can inhibit others in their freedom of speech in public debate, and in that way weaken democracy. Moreover, hate speech can keep prejudices alive, deprive people of their dignity and cause fear and alarm in the groups it targets. Recurrent hate speech targeting selected groups can serve to legitimise harassment and discrimination, and ultimately violence towards individual members of these groups (Norwegian Equality and Anti-Discrimination Ombud 2015; Nilsen 2014).

The purpose of this report is to review research that sheds light on the following subjects:

- the extent of online hate speech
- which groups in society are particularly subjected to online hate speech
- who produces hate speech, and what motivates them

The report does not address whether or how hate speech should be prohibited by law. The purpose is rather to collate research on the nature and extent of hate speech in Norway and internationally and factors to account for the prevalence

of this phenomenon. In addition, we present updated figures on the extent of hate speech in social media in Norway, based on questionnaire-based surveys conducted in 2013, 2015 and 2016.

This report is part of a series of three reports on hate speech.

Hate speech, report 2. Research on hate and discrimination (Eggebø & Stubberud 2016) reviews research that highlights:
- the relationship between hate speech and discrimination, bullying and violence
- research on the consequences for groups subjected to the above, and for society.

Hate speech, report 3. The boundary between freedom of speech and criminal law protection against hate speech (Wessel-Aas, Fladmoe & Nadim 2016) reviews:
- how hate speech is defined and regulated internationally, in national legislation and relevant legal practice

In addition, the report elucidates ongoing discussions of the boundary between freedom of speech and protection against hate speech, based on existing research.

1.1 What is hate speech?

There is no shared and agreed upon definition term hate speech. However, hate speech is commonly understood as persecutory, degrading or discriminatory speech directed at a specific *minority group* or an individual's (assumed) group identity. Thus, to be considered hate speech, the *grounds* it targets are of significance. The above definition particularly comprises minority groups, who are assumed to be especially vulnerable.

The Norwegian legal provision against hate speech (Section 185 of the Norwegian Penal Code, formerly Section 135a) provides legal protection against discrimination or hate speech on grounds of: a) skin colour or national or ethnic origin, b) religion or life stance, c) homosexual orientation or d) disability. In other words, hate speech that is equally offensive in nature but based on other grounds will not be comprised by Section 185 of the Norwegian Penal Code. It is worth noting that the legal provision provides protection for all grounds comprised by the Norwegian anti-discrimination law, with the exception of gender, gender identity and gender expression. This means that Norwegian law on hate speech does not cover hate speech motivated by hatred of, for example, women or transgender people, as groups.[2]

2 Report 3 (Wessel-Aas et al. 2016), however, demonstrates that hateful speech that does not target the groups protected by Section 185 of the Norwegian Penal Code, may still be prohibited under other provisions in the Code.

Section 185 of the Norwegian Penal Code, which is devoted specifically to hate speech (Norwegian: "*Hatefulle ytringer*"), further defines hate speech as expressions that are "threatening or insulting anyone, or inciting hatred or persecution of or contempt for anyone" because of the grounds specified above.

In public and political debate, the term "hate speech" is, however, often used in a broader sense than what ensues from the Norwegian Penal Code. The term is, for example, used in reference to speech expressing contempt for humanity, intolerance or aggression, and includes a wide range of phenomena, from online bullying and aggressive and intolerant outbursts in public debate, to racism and criminal threats against individuals (Sunde 2013: 42).

In seeking to understand and study the phenomenon of hate speech, it may be useful to apply a broader definition of the term than what ensues from Section 185 of the Norwegian Penal Code. The justification for this is in part methodological, in that it is extremely challenging to delimit hate speech so that it applies solely to instances covered by the Norwegian Penal Code. Few people are aware of what is legally defined as "hate speech", and the line between criminal hate speech and other similar, but non-criminal speech will in many cases be blurred. But a broader definition of the term also has a substantive rationale. Speech that is not comprised by the law can still have adverse consequences for individuals and society at large. In order to gain a sense of what motivates the perpetrators of hate speech, what it is like to be at the receiving end of hate speech, and what the consequences might be, the legal definition of hate speech is unlikely to be adequate (cf. Norwegian Equality and Anti-Discrimination Ombud, 2015). Although it might be useful to operate with a broader understanding in studies of hate speech, it is important to emphasise that this is not an argument in itself for amending the prevailing Norwegian legal definition of the phenomenon. Protection against hate speech must always be weighed up against respect for freedom of speech (for a more detailed discussion, see Report 3 by Wessel-Aas et al. 2016).

One example of a broader approach to hate speech was proposed by the Norwegian Equality and Anti-Discrimination Ombud. The principal argument in this proposal is that in the interests of prevention, it is necessary to include both legal and illegal hate speech. The Ombud also points out that hate speech that is not governed by law can have adverse consequences and that the boundary between lawful and unlawful speech is in any case unclear. On this basis, the Ombud proposes an extended definition of the phenomenon:

> Hate speech is degrading, threatening, harassing or stigmatising speech which affects an individual's or a group's dignity, reputation and status in society by means of linguistic and visual effects that promote negative feelings, attitudes and perceptions based on characteristics such as ethnicity, religion, gender, disability, sexual orientation, gender expression, gender identity and age. (Norwegian Equality and Anti-Discrimination Ombud, 2015: 12).

This definition takes account of both the *intention* behind the hate speech and the *effect* it may have on the individual(s) it targets. Unlike the Norwegian Penal Code, there is no condition that it must constitute wilful or gross negligence or that the speech must be conveyed in the presence of others. Moreover, the definition covers more grounds than Section 185, and the list is not exhaustive. Among other characteristics, gender, gender expression, gender identity and age are also named.

Hate speech, including that classed as non-criminal, tends to be founded on negative stereotypes, prejudices and stigma. The object is to demonstrate differences between groups. Hate speech is not necessarily motivated by a strong sense of hatred (Ask, Svendsen & Karlstrøm 2016; Erjavec & Kovačič 2012), but tends to embody an exclusion rhetoric, and plays on notions of inherent hierarchies, irrational fear and contempt of individuals and groups who are regarded as different (Norwegian Equality and Anti-Discrimination Ombud 2015; Nilsen 2014; Perry 2001).

In this report, we will be discussing research that concerns criminal hate speech, non-criminal hate speech and speech that may be *perceived* as hate speech. As we will see, the lines between these phenomena are often blurred.

1.2 Hate speech versus other offensive speech

Many different terms are used to refer to the negative sides of online discourse. Hate speech is one aspect of this, alongside phenomena such as online threats, online harassment, trolling and flaming[3] (see for example Hagen 2015: 35 for a list of terms). A number of studies have emerged of distressing experiences on the internet, of trolling, online harassment and the like. However, this research rarely examines the *grounds* for the offensive comments, meaning what the comments are targeting. As mentioned, a key aspect of hate speech is that it targets specific and (presumed) at-risk, group identities or minority statuses.

3 Deliberate, hostile argumentativeness, often using offensive, crude and vulgar language, insults and threats.

Without information about what offensive comments are targeting, we cannot determine whether we are dealing with hate speech or other forms of online hate.

For hate crime, however, there is a relatively extensive body of literature. This tends specifically to examine criminal offences, violence or threats motivated by "hate" towards a minority group (see Hall 2013 for a discussion of the various definitions of hate crime). Hate speech is one form of hate crime, but the term comprises a range of other types of incidents such as physical harassment and violence. Previous studies rarely distinguish between verbal and physical forms of hate crime. Moreover, hate crime usually denotes criminal offences. As such, only the most extreme hate speech can typically be brought under the definition of hate crime.

Many of the studies discussed in this report examined either negative experiences online or hate crime. This means that many of the findings do not strictly refer directly to hate speech. We will be reserving use of the term "hate speech" for offensive, discriminatory or degrading expressions linked to a group identity in line with the Ombud's expanded definition, and will otherwise be adopting terms from the cited research (for example "distressing experiences" and "harassment" online).

1.3 A note on the literature search and data sources

This literature review is based on existing research literature on hate speech. We conducted systematic literature searches in internationally-oriented reference databases under Web of Science and ProQuest, which includes Social Sciences Citation Index, Sociological Abstracts and several others. Searches were made in the Danish and Swedish union catalogues (DANBIB and LIBRIS) and in the Norwegian article index NORART in addition to Oria (BIBSYS interface). These databases cover the main journals for international social science publications in addition to books and chapters in books. The primary publications in Norwegian, Danish and Swedish were also identified by searching the national union catalogues.

The search terms were "hatefulle ytringer", "hat prat", "hatytringer" and "hate speech". When this yielded only very limited results, the decision was made to include the search words "hate crime", "online hate" and "online extremism".

A manual review of the search results revealed that these included a great deal of literature of no direct relevance to the present report. The research literature on hate speech has focused largely on normative issues relating to the tension between freedom of speech and protection against hate speech, and much of the literature is characterised by philosophical, normative or legalistic approaches. There is considerably less empirical research on hate speech as a phenomenon, its scale and its nature.

Due to the challenges posed by the systematic searches, we supplemented them with manual searches. We have, for example, reviewed the reference lists of relevant publications (so-called snowball method), run keyword searches in Google and Google Scholar and sent specific requests for information to Nordic peers in the field of research concerned. The latter came about as a result of a remarkable lack of Danish and Swedish research papers in the systematic searches. We were also obliged to broaden the scope of our literature search beyond that dealing specifically with hate speech. Firstly, we included a proportion of the relatively comprehensive literature on hate crime. Secondly, we included key studies concerning bullying, harassment and other negative online experiences. This means that a number of the studies we refer to describe negative experiences that were not necessarily targeting a specific (minority) group identity. Consequently, we cannot determine if the reported experiences could be characterised as hate speech.

For practical reasons, we chose to limit the systematic review to research that as directly as possible addresses online hate speech. Because the scale of hate speech was focal for the systematic review, our primary emphasis was on quantitative studies. The topic exists in a cross-field between many other fields of research such as hate crime, discrimination, bullying, extremism and violence, and studies in these fields may be relevant in elucidating the nature and extent of hate speech. This more indirect approach to the field is, however, beyond the remit of this project, and was therefore not pursued systematically.

In addition to the systematic review, the report also contains an independent analysis chapter devoted to the extent of hate speech in Norway. Here we present the results of a number of surveys on experiences of hate speech and other offensive speech. These surveys were conducted by the Institute for Social Research in the period 2013–2016, and a number of the figures have not been published previously. This material makes it possible to determine what proportion of the Norwegian population has been subjected to what they perceive as hate speech in social media. In addition, the surveys allow us to

compare the majority population's experiences with the experiences of people of immigrant descent from Eastern Europe, Asia and Africa, and with journalists.

1.4 The structure of the report

The remainder of the report is divided into four chapters. In the next chapter, we review existing research on the extent of online hate speech; which groups are particularly at risk, and the platforms used. Key sources include statistics based on Norwegian police crime-report records and large-scale social science questionnaire-based surveys of self-reported experiences. Chapter Three presents new analyses of the extent of hate speech in Norway based on surveys conducted in the period 2013–2016. Taken together, these surveys make it possible to distinguish between majority-population experiences and the experiences of journalists and people of immigrant descent from Eastern Europe, Asia and Africa. In Chapter Four, we review existing research on the perpetrators of hate speech. What characterises them, and what are their motives? Finally, the present report summarises the status of research in the field and discusses research needs.

2 Research on the extent of online hate speech

In this section, we review sources, which in various ways can provide us with a picture of the extent of hate speech. We start by looking at how many people report hate speech and hate crime to the police. We then review various surveys that have sought to map the extent by inquiring into people's experiences of the phenomenon; essentially by means of questionnaires. As stated earlier, scarcely any of these surveys have operated with a restricted definition of "hate speech". Instead, they tend to measure distressing experiences or harassment online, often without information about whether or not the targets were specific group identities. Nevertheless, these surveys still provide relevant data on the extent of *potential* online hate speech.

Given that questionnaire-based surveys are conducted with representative samples, one advantage of the method is that it makes it possible to apply the results generally to the proportion of the population studied. In other words, the surveys provide relatively reliable data on the extent of hate speech and other offensive and negative statements in the population, and on which groups are most subjected to such speech.

However, studies based on questionnaires also have some obvious limitations. Firstly, because they survey *subjective experiences* of hate speech. We cannot tell how serious the different experiences have been, and two different respondents in a single survey may perceive the same speech differently. To some extent, this can be avoided by asking questions about experiences of specific incidents, as the problem is greatest in surveys where the wording of questions is relatively general. In other words, questionnaire-based surveys cannot provide an objective estimate of the extent of hate speech, but rather an estimate of what the respondents *perceive as* hate speech or offensive speech, and how these experiences vary between different segments of the study population.

Secondly, a number of the surveys reviewed in this report are based on population-representative samples, or samples in which all segments of the population are in theory equally represented. This means that in absolute figures there will be relatively few respondents who represent the minority groups afforded special protection by Norwegian law against hate speech. This problem is compounded

by the fact that some minority groups are often under-represented in population-representative surveys. We know, for example, that certain immigrant groups account for a distinctly lower response rate than the general population (Djuve, Gulløy, Kavli & Berglund 2009), which may cause the extent of racist hate speech, for example, to be under-estimated. Under the section on research needs in the concluding chapter, we discuss methods for overcoming such problems in future research.

2.1 Crime-reporting statistics

At the national level in Norway, we only have access to data on reported hate crimes generally. The number of hate crime cases in police records is low in Norway, also in relation to the overall crime rate (Norwegian National Police Directorate 2015):

- In 2014, 223 reports were made in which hate crime was recorded as the motive.
- In the period 2010–2014, the hate crime reporting rate was between 216 and 307 per annum.[4]
- The largest category by far for the hate motive is "race/ethnicity" (Norwegian National Police Directorate 2015).
- For purposes of comparison, Denmark recorded 139 hate crime reports in 2011, while Sweden recorded 5490 in the same year (Oslo Police District 2013). The substantial differences between the Scandinavian countries are attributable to differing recording procedures, but presumably also a difference in enforcement, competence and awareness of hate crime within the police and civil society (Oslo Police District 2013).

Oslo Police District has, however, devoted special attention to hate crime and has published more detailed reporting figures for its own district. These figures also include information about what type of group identity was targeted or the grounds on which the hate crime was based (Oslo Police District 2016):

- In Oslo in 2015, there were 41 reports of hate speech, in addition to 22 reports of threats that were defined as hate crime.
- Hate speech is one of the largest categories of crime under hate crime, followed by "common assault/battery".
- It is not possible to separate online hate speech out as a distinct category in the crime reporting statistics, but in 20 of the cases the crime scene was recorded as social media, e-mail or telephone.

4 Changes in the procedures for recording such reports mean that the figures in these years are not directly comparable.

- Hate crime cases in which “media” was recorded as the crime scene break down evenly by ethnicity and religion as the grounds, while there were very few reports with sexual orientation as the grounds.
- The records show a marked increase in cases of reported hate speech in Oslo, from 15 in 2014 to 41 in the following year. Prior to 2014, hardly any hate speech reports were filed. This increase does not necessarily reflect a genuine increase in the incidence of hate speech, but more likely the increased efforts and vigilance surrounding the phenomenon within the Oslo Police District.

Based on the information available, it is difficult to find the exact number of reports of cases of online hate speech. However, nationwide in Norway, we find relatively few reports of hate crime generally, and in Oslo, in the peak year 2015, there were just the 41 cases of hate speech reported. The police believe that the reporting statistics reflect ”double” underreporting. Firstly, because many members of the public refrain from reporting this type of incident, and secondly because for various reasons, a number of cases are not recorded as hate crimes in police records (Norwegian Directorate of Police 2015). This means that the crime reporting statistics are not ideally suited to providing a comprehensive picture of the extent of online hate speech in Norway.

2.2 The extent of experiences of hate speech and offensive experiences online

In the following, we review key studies which in various ways shed light on the extent of online hate speech and offensive speech both in Norway and internationally. All the studies involve questionnaire-based surveys conducted among representative sample of the entire, or certain segments of, the population in the countries where the studies were conducted.

Hate crime

First, we will be reviewing two studies revealing the extent of experiences of *hate crime*. This is thus a broader concept than hate speech in that it also includes other forms of actions such as violence motivated by “hate”, or negative attitudes to certain minority groups.

In the Norwegian Police population surveys (Norwegian Directorate of Police 2012, 2016), representative samples of the Norwegian adult population were asked if they had been subjected to different types of crime, including whether they had been the targets of “hate crime (such as violence, threats or other crime)”.

- The 2012 survey showed that young adults (18–34-year-olds) generally, and young adults of immigrant descent from non-Western countries especially, were the most frequent targets of hate crime. 3 per cent of those who were not of immigrant descent and 8 per cent of those of immigrant descent from non-Western countries in this age-group responded that they had been the targets of "hate crime such as violence, threats or other crimes on grounds of belief, skin colour, ethnicity, nationality or sexual orientation". The corresponding proportion of the entire sample was 1 per cent (Norwegian Directorate of the Police 2012).
- No similar survey was conducted in 2015 of persons of immigrant descent, but a total of 1.7 per cent of respondents stated that they had been the targets of hate crime. As such, there was no substantial change since the previous survey in terms of experiences of hate crime in the general population (Norwegian Directorate of Police 2016: 28).

A report on hate crime in Denmark (COWI 2015) includes several circumstances under the term hate crime (including harassment), in addition to the specific grounds the crime was directed at. This Danish survey produces somewhat higher figures than the Norwegian Police population surveys. The report is based on data from 2014.

- A total of 3 per cent stated that they had definitely – and 10 per cent that they had possibly – been the targets of hate crime within the past year (a total of 13 per cent).
- Harassment was the commonest form of hate crime (37 per cent of instances), which implies that around 5 per cent of the Danish population definitely or possibly experienced harassment within the last year.
- The commonest grounds for hate crime were gender, followed by social status, age and political or other beliefs. Further down the list were ethnicity or skin colour (3.6 per cent definitely or possibly), religion and beliefs (2.4 per cent), disability (1.6 per cent), sexual orientation (1.0 per cent) and gender identity (0.6 per cent).

It is important to interpret these results in the light of the fact that the proportion of persons *potentially* targeted by the various forms of hate crime and harassment varies. "Everyone" is potentially exposed to harassment based on gender, while far fewer will be subjected to harassment based on sexual orientation or disability. At the same time, harassment or verbal abuse concerning sexual orientation are not necessarily only directed at sexual minorities. We will be returning to which groups are particularly vulnerable to online hate speech and harassment later in this chapter.

Online hate speech

One of the very few studies we found that seeks to measure the extent of hate speechis a comparative study from 2013–2014, which compares experiences of online hate speech among young adults (15–30-year-olds) in the US, UK, Germany and Finland (Hawdon, Oksanen & Räsänen 2015).

- This study indicates that the proportion that *witnessed* online hate speech within the last three months varied from 31 per cent in Germany to 53 per cent in the US.
- The proportion that had personally been *targeted* by hate speech or degrading statements varied from 4 per cent in Germany to 16 per cent in the US. In the UK and Finland, around 10 per cent responded that they had been targeted by such speech.

The authors suggest that the high incidence of hate speech in the US may be attributable to the strong constitutional protection of freedom of speech and reluctance to prohibit hate speech in law. The US also has many and prominent organised hate groups that operate online (Hawdon et al. 2015: 34). We were unable to find comparable surveys directly studying hate speech in a Norwegian context, but in Chapter 3 we present analyses of studies that serve to shed light on the extent of hate speech in Norway.

Offensive or derogatory comments

A Norwegian study surveys the frequency with which people receive *offensive or derogatory comments* after expressing themselves publicly.

As part of a research project funded by the Fritt Ord Foundation entitled *Status for ytringsfriheten i Norge* (*Status of Freedom of Speech in Norway*), nationally representative questionnaire-based surveys were conducted in 2013 among individuals in the majority population and among individuals of immigrant descent from Asia, Africa and Eastern Europe (Staksrud et al. 2014). Those who had participated in a discussion or commented publicly in at least one arena (including social media) were asked if they had experienced receiving offensive or derogatory comments afterwards.

- 31 per cent of the majority population responded that they had had such negative experiences, while 28 per cent of individuals of immigrant descent reported the same. This difference is not statistically significant.
- Most had experienced receiving offensive comments on several occasions: Around half of both samples responded that they had had negative experiences of this nature on between two and ten occasions, while between 10 and 15 per cent responded that they had had such experiences more frequently.

- Although the differences between the majority population and individuals of immigrant descent were minor, there were substantial differences in what the comments were most typically directed at:
  - the markedly most frequent grounds for offensive comments aimed at individuals in the majority population were the content of their contentions and political opinions.
  - The most frequent grounds among individuals of immigrant descent were religion, ethnicity, nationality and skin colour.

We will be presenting more thorough analyses of this data material in the next chapter.

Offensive experiences online

An American study of adult internet users from 2014 (Pew Research Center 2014) surveys experiences of different types of offensive online encounters. The study shows that it is very common to have witnessed someone being harassed online. A relatively large number had also personally experienced various forms of online harassment.

- 73 per cent had *witnessed* someone else being harassed in some way online, but this also included a number of less severe forms of harassment. 60 per cent had seen someone being called offensive names, 25 per cent had seen someone being physically threatened, 24 per cent had seen someone being harassed for a sustained period of time and 19 per cent had witnessed someone being sexually harassed.
- A total of 40 per cent responded that they had personally experienced one of the forms of harassment they were asked about. 22 per cent had experienced "less severe" types of harassment (name calling or embarrassment), while 18 per cent had fallen victim to more severe forms of harassment (physical threats, harassment over a sustained period of time, stalking and sexual harassment).
- 27 per cent of the respondents had experienced name calling, 8 per cent had been the targets of physical threats and between 6 and 7 per cent had been harassed over a sustained period of time and/or been sexually harassed.

This survey, however, provides no information as to whether the experiences of harassment were linked to specific group identities.

Table 1 provides an overview of the three surveys of the extent of experiences of different forms of offensive speech and experiences.

Table 1. Surveys of experiences of different forms of hate speech or offensive experiences online

Source	Conducted (country/year)	What is surveyed?	Extent in per cent	Target group
Hawdon, Oksanen & Räsänen (2015)	Finland 2013 USA 2013 UK 2014 Germany 2014	Experienced being the target of hate speech or derogatory materials online	10 16 12 4	15–30-year-olds
Pew Research Center (2014)	USA 2014	Name calling Threatened physically Harassed over a sustained period of time Sexually harassed	27 8 7 6	Age 18+
Status of Freedom of Speech in Norway (Staksrud et al. 2014)	Norway 2013	Experienced receiving offensive or derogatory comments after participating in a discussion or stating an opinion publicly	31 28	Majority population age 16+ Individuals of immigrant origin from Asia, Africa and Eastern Europe, age 16+

The table demonstrates that different surveys provide different pictures of the extent of negative experiences online and hate crime. The more severe the phenomenon measured by the surveys, the smaller the extent. While it is not possible to extrapolate an "objective" figure for the extent of online hate speech from these surveys, they still indicate that harassment – which may range from insults to more threatening behaviour – is a relatively common part of life on the internet. In the next chapter, we supplement this overview with updated analyses of the extent of hate speech and other offensive speech in Norway.

2.3 The extent of cyberbullying among children and adolescents

A number of studies exist on cyberbullying among children and adolescents. These do not measure hate speech specifically because they hold no data on whether or not the bullying targeted a group identity. While bullying generally targets individuals, as stated, hate speech targets group identities or group affinities. These studies do, however, provide an impression of the extent of child and youth experiences of distressing online experiences.

Youth in Norway survey

The *Ung i Norge* Norwegian nationwide questionnaire-based survey of adolescents (13–15-year-olds) has been conducted several times since 1992 (NOVA 2015). The questionnaires are completed during school hours and yield very high response rates. A summary of the findings of the surveys from 2012 to 2014 show that relatively many adolescents have experienced being abused or threatened online, but that few of them have experienced this regularly.

- A total of around one quarter of the adolescents responded that they had experienced abuse or being threatened online.
- 6 per cent responded that they had been subjected to abuse and threats by other teens over the internet or by mobile phone at least once a month. 3 per cent responded that they suffered abuse at least every fortnight.

Norwegian Children and Media Survey

The Children and Media Survey (*Barn og medier-undersøkelsen*) (Norwegian Media Authority 2016) is a survey of children aged 9–16, and has been conducted several times since 2008. It reveals the same general patterns as the *Ung i Norge* survey of adolescents: Relatively many have experienced someone being nasty towards them or bullying them online, but few of them experience this on a regular basis. It is also common to have witnessed cyberbullying and relatively many have also witnessed someone being threatened.

- In the survey from 2016, 25 per cent responded that they had experienced someone being nasty towards them or bullying them online. 7 per cent responded that they had been subjected to this every month or more frequently.
- 16 per cent of children responded that they had experienced being threatened online. 5 per cent experienced this every month or more frequently.

The period 2014–2016 showed a strong increase in the proportion who experienced someone being nasty towards them, bullying or threatening them.[5]

The 2016 survey also found that:

- 41 per cent have seen someone be nasty towards or bully others online. 16 per cent experienced this every month or more frequently.
- 25 per cent witnessed threats online, 9 per cent every month or more frequently.
- 10 per cent have bullied or been nasty towards someone online, 2 per cent being so every month or more frequently.

5 According to the Norwegian Media Authority, some of the increase is attributable to the fact that more 16-year-olds were included in the sample in 2016, and that bullying and threats are more prevalent among teenagers. However, even taking into account the inclusion of more of the higher age-group in the sample, the scale of bullying and threats had increased markedly from 2014. (Personal communication with the Norwegian Media authority).

EU Kids Online

EU Kids Online (Livingstone, Haddon, Görzig & Ólafsson 2011) is an international survey of children and adolescents aged 9–16 in which 25 countries participated. The findings from Norway showed that:

- 8 per cent had been bullied online within the preceding year.
- The most common experience was receiving nasty or hurtful messages or that nasty or hurtful things were spread on the internet.
- Norway had the fifth-highest figure in the survey for the extent of online bullying.

The main findings from the three surveys are summarised in table 2.

Table 2. Surveys on the extent of cyberbullying among children and adolescents

Source	Survey year	What is surveyed?	Extent in per cent	Target group
Ung i Norge (NOVA 2015)	2012–2014	Abuse and threats ...monthly or more frequently ...at least every fortnight	 6 3	13–15 år
Norwegian Children and Media Survey (Norwegian Media Authority 2016)	2016	Someone was nasty towards you or bullied you ... over the last year ... monthly or more frequently Threatened ... over the last year ... monthly or more frequently	 25 7 16 5	9–16 år
EU Kids Online (Livingstone et al.) 2011)	2010	Been bullied over last year	8	9–16 år

The two Norwegian surveys of child and adolescent experiences paint a relatively similar picture of the extent of cyberbullying. Both the Norwegian Media Authority and NOVA's youth data scheme find that 6–7 per cent suffer abuse/experienced someone being nasty/have been bullied at least every monthly. The Norwegian Media Authority's survey shows that 25 per cent were bullied, while 16 per cent were threatened *within the preceding year.* The EU Kids Online multinational research network reports substantially lower figures for cyberbullying last year (8 per cent), which is presumably attributable to the fact that this study restricts the question to cyberbullying. The Norwegian Media Authority includes less serious incidents in which someone was "nasty".

2.4 What groups are especially at risk?

The questionnaires we used to gain a picture of the extent of hate speech also give some indication of who is especially at risk. However, it is important to distinguish between who is at risk and which grounds the hate speech targets. Several of the surveys show that the differences between groups are not always that great when it comes to the proportion who had distressing experiences online, but that there may still be great variation in what the negative statements targeted.

In addition, it is important to note that speech targeting specific grounds may not necessarily be directed exclusively at members of the group in question. For example, Norwegian words like "homo" ('gay') and "jøde" ('Jew') are used as general terms of abuse (Helseth 2007; Hoffmann, Kopperud & Moe 2012). Although the speech is not aimed directly at members of the minority group in question, it may still impact that group if its members experience such terms used as insults. In other words, individuals who identify with certain groups may be exposed to hate speech even in instances where they are not the direct recipients of the abuse. This is also emphasised by Section 185 of the Norwegian Penal Code in that hate speech is unlawful both when conveyed publicly and when conveyed in the presence of others (see Report 3 (Wessel-Aas et al. 2016) for further discussion).

The surveys we have reviewed are not designed to determine who is subjected to online hate speech and hate crime. For example, several of the grounds in the definition of hate speech, such as disability, are not included. Below, we summarise research on the different grounds that emerge from general studies of negative online experiences. In this report, we have not systematically reviewed the research conducted on the experiences of different groups (but see Report 2 (Eggebø & Stubberud 2016), which presents a systematic review of the literature on the different groups). In addition to the groups protected by Section 185 on hate speech in the Norwegian penal code (ethnicity, religion, sexual orientation and disability), we review variations linked to gender and age. Finally, we present some research on one exposed professional group for which data is available, namely journalists.

Age

Young adults appear to be an especially exposed group.

- In the survey by Pew Research Centre, 70 per cent of 18–24-year-olds responded that they had experienced a form of harassment (including less severe forms) online. Almost a quarter had received physical threats (Pew

Research Center 2014: 14). For the sample as a whole, the corresponding figures were respectively 40 and 8 per cent.

- Cyberbullying is more prevalent among adolescents than young children. Older children have more frequently experienced being bullied, and have more frequently witnessed bullying (Livingstone et al. 2011; Norwegian Media Authority 2016).

Gender

The surveys generally find only minor gender differences when it comes to extent overall, but that men and women experience different types of online hate and harassment.

- While men primarily experience offensive and derogatory comments aimed at the content of their contentions, women experience that much of the online hate targets gender and appearance (Staksrud et al. 2014: 41).
- The survey by Pew Research Center shows that men overall experience harassment more frequently than women, but there are distinct differences in the type of harassment experienced by men and women. Men tend more to experience name-calling and receiving physical threats, while women are more the targets of sexual harassment and stalking (Pew Research Center 2014: 5).
- The comparative study by Hawdon et al.found significant differences between the countries regarding the extent to which people had witnessed hate targeting gender. While 44 per cent of the hate observed in the US and UK was linked to gender, the proportion of hate linked to gender was almost half as great in Finland and Germany (Hawdon et al. 2015: 34). Aside from the UK, where gender represented the third-most frequent grounds for hate speech, gender was far down on the list of different grounds specified in the survey.
- The studies of cyberbullying likewise find no systematic gender differences, but that boys are at higher risk of being subjected to threats.
  - The most recent Ungdata report on youth in Norway from NOVA found that girls were more at risk than boys, but among those who frequently suffered abuse or bullying online (at least every fortnight), there was no gender difference (NOVA 2015: 103).
  - In the latest Children and Media Survey by the Norwegian Media Authority, no systematic gender differences were found for bullying. The survey did, however, find that more boys than girls had been subjected to threats. A total of 31 per cent of the boys and 18 per cent of the girls aged 15–16 responded that they had been threatened on the internet, gaming sites or by mobile phone (Norwegian Media Authority 2016: 67).

In recent years, much attention has been given to how women as a group are subjected to online threats, hate and harassment, and the implications of this for freedom of speech. Report 2 (Eggebø & Stubberud 2016), however, point to the fact that research in this field is relatively limited.

Ethnicity, religion etc.

Ethnicity, skin colour and religion feature among the most frequent grounds for hate speech in several surveys.

- www.hatebase.org is a global database of instances of hate speech. Their summary statistics show that the vast majority of instances of hate speech in their database targeted individuals, based on ethnicity and nationality (Hatebase 2016).
- Ethnicity and religion are also among the leading grounds for the hate which the respondents report they witnessed, in the Hawdon et al. (2015: 34) international survey.
- Data from the 'Status of Freedom of Speech in Norway' survey showed that there were generally only minor differences between the majority population and individuals of non-Western descent when it came to the proportion that had received offensive or derogatory comments online. The ethnic minority population, however, reported a higher incidence of comments targeting skin colour, religion and nationality, that is, grounds protected by the Norwegian Penal Code (Midtbøen & Steen-Johnsen 2016; Staksrud et al. 2014). We will be examining these data in more detail in the next chapter.

Report 2 (Eggebø & Stubberud 2016) indicate that extremely limited research has been conducted on hate speech, discrimination and bullying targeting indigenous peoples and the Norwegian national minorities such as the Sami people, Kven people, Forest Finns and indigenous travellers/Romani people, while a few studies have addressed the experiences of Norwegian Jews, where the prevalence of anti-Semitism in the general population is examined.

Sexual orientation

- In the four countries included in the Hawdon et al. comparative study, sexual orientation ranks first or second among the grounds people most commonly report they have witnessed (Hawdon et al. 2015: 34).
- Report 2 (Eggebø & Stubberud 2016) reviews research showing that bisexual, homosexual and lesbian pupils in Norway are significantly more subjected to bullying by mobile phone or online than other pupils. Young

homosexual males especially are subjected to bullying targeting their sexual orientation (Roland & Auestad 2009: 34–35).
- Among adult lesbians and homosexuals in Norway, around two in ten have experienced negative comments or negative behaviour at the workplace within the last five years on the grounds of their sexual orientation. The proportion was somewhat lower among bisexual women and men (Anderssen & Malterud 2013: 92).

There is no information on sexual orientation in the other surveys we have reviewed here.

Disability

- Between 13 and 18 per cent of online hate observed in the four countries included in the Hawdon et al. study are linked to physical disability (Hawdon et al. 2015: 34). This is one of the grounds respondents least-frequently report having witnessed as a hate target.
- At the same time, there is much to indicate that hate crime perpetrated against people with disabilities is significantly underreported (Digranes 2016; Hall 2013).
- Statistics Norway's figures show that people with disabilities are at three times the risk of being subjected to physical assault or threats of violence than the general population (Ramm 2010: 62).

Report 2 (Eggebø & Stubberud 2016) indicates that research on hate crime targeting people with disabilities is only an emergent field of research with relatively scant existing research. There are at present no Norwegian studies of hate crime or hate speech targeting this group. The Norwegian Directorate for Children, Youth and Family Affairs has, however, commissioned a study of hate speech targeting people with disabilities (Olsen, Vedeler, Eriksen & Elvegård 2016).[6]

Online hate and threats against journalists

Certain professional groups may be especially exposed to hate speech. Scarce attention has been given to hate speech in working life, but separate studies have been conducted on journalists' experiences of harassment and threats in all three Scandinavian countries. Due to their public role, journalists are likely to be especially subjected to hate speech and other offensive statements. The

6 This study will be conducted by the Nordland Research Institute and the NTNU Samfunnsforskning social sciences research institute and was scheduled for completion by autumn 2016, but had not been published when this report went to press.

studies in question do not examine online hate speech specifically, but show that journalists as a group are vulnerable to harassment and that certain case areas elicit more hate reactions than others.

- In Norway, just over 40 per cent of journalists responded that they have been subjected to harassment, sexual harassment, persecution or obscenities over the last five years, while a quarter have received threats (Hagen 2015: 13).
- In Sweden and Denmark, between 40 and 45 per cent of the journalists respond that they have been subjected to harassment (Mølster 2015: 58).
- The extent to which journalists are subjected to harassment and threats is linked to the types of topics they cover. Topics such as immigration, political conflicts and gender equality are the topics that trigger most hate reactions (Hagen 2015; Nilsson 2015).

2.5 Which platforms are used?

The platforms that form the digital crime scene for online hate speech and hate crime change rapidly. New platforms arise and become immensely popular within a short space of time, while others are abandoned by the users after a while. The manner in which the different platforms are used also changes over time. This means that the studies described below only provide a snapshot of the lay of the land at the time when the studies were conducted.

- Foxmann and Wolf (2013) describe different ways in which the internet is used for spreading and inciting hate, and highlight platforms such as social media, online games, websites promoting extremist groups and cloaked websites that are ostensibly impartial and factual, but actually consist of hate propaganda. The American research literature on hate crime is particularly taken up with online hate sites and hate groups. However, the hate content is not necessarily read by the groups targeted by the hate.
- Pew Research Center (2014)[7] found that 66 per cent of those who had experienced online harassment had done so on social networks, while 22 per cent responded that the negative experience took place in a website comment field. In 16 per cent of those who experienced harassment, the platform was e-mail or online games.
  - Women and young adults are especially at risk on social media, while men have a higher probability of experiencing harassment in online games and comment fields. Older internet users (50+) have a higher probability than

7 As stated, the survey contains no information about the grounds for the harassment. This means we have no way of telling if it comes under the category of hate speech.

others of experiencing harassment by e-mail (Pew Research Center 2014: 24–25).

- Hawdon et al. (2015) also studied platforms for hate speech. They found Facebook to be the most frequent platform on which people had *witnessed* hate, with YouTube in second place. Twitter was high on the list in the US and UK, but was a less common channel in Finland and Germany. Besides Facebook and YouTube being in first and second place in all four countries, there were major differences in which platforms people reported as being digital crime scenes for hate speech.
- In the EU Kids Online 2011 study, it emerged that social networks and instant messaging were the commonest platforms for cyberbullying. E-mail, online games and chat rooms were less common arenas for bullying, which was explained by the fact that these arenas were generally used less (Livingstone et al. 2011: 63).
- In a Norwegian survey of 11–12-year-olds from 2015, Snapchat was by far the highest on the list of platforms on which children had experienced nastiness, bullying or threats. This is a photo and video sharing app in which the messages are deleted within ten seconds. Other services that were the scene of distressing experiences included Instagram (photo sharing app), Moviestar Planet (a mixture of online community and gaming for children) and, to a lesser extent, Facebook (Aftenposten 2016).[8]

2.6 Concern about being subjected to hate speech

In the foregoing, we looked at research on the extent of hate speech; who is targeted and where it occurs. But is this a phenomenon people are aware of, and which they are concerned about being subjected to? Three Scandinavian studies take different approaches in surveying the extent to which people are concerned about being subjected to distressing experiences on line.

One of the questions in the Norwegian Police population survey of 2016 asked people about the extent to which they were concerned about being targeted by different types of crime (Norwegian Directorate of Police 2016). This survey revealed that:

- Around one in ten was concerned about being targeted by online harassment and hate crime to a moderate or very great extent. This corresponded with the level of concern about disturbances of the peace and violence or threats of violence.

8 We have tried to gain access to this survey, but without luck.

In the Danish hate crime survey (COWI 2015), those *who had experienced hate crime* were asked if they felt they were at risk of crimes targeting different grounds.

- A total of 9 per cent of those who had experienced hate crime within the last year felt they were at risk on grounds of the various identity categories.
- Women, adolescents and people with a low income/short education saw themselves as being the most at risk of hate crime.

A Swedish study on social media surveyed the perceived risk of being the target of online hate and personal attacks on social media (Ghersetti 2015). The study indicated the following:

- Two in three completely or partly agreed that they risked being subjected to online hate and personal attacks on social media.
- Women were somewhat more concerned than men, and the young were more concerned than older people. There was little difference between daily users and infrequent users of social media.

The three studies paint very different pictures of the population's level of concern about being at risk of distressing experiences online, and the differences presumably largely reflect the fact that the questions were posed in different contexts. In the Norwegian Police Directorate's survey, respondents are asked about their concern about online harassment and hate crime in the context of their concern about a whole series of criminal offences, some of which were very serious. This can lead the respondents to interpret the questions about online harassment as concerning serious and potentially criminal harassment. The Danish survey, however, does not distinguish between different forms of hate crime and does not report on concern among those who have not experienced hate crime within the last year. This means that we do not know exactly what type of crime the Danes were concerned about, or what proportion of the population as a whole is concerned. The Swedish survey, for its part, deals with social media and the wide degree of concern it reports may reflect an elevated awareness of online risks in the general population.

In the next chapter, we present previously unpublished findings from recently conducted surveys in Norway on experiences of hate speech and other distressing statements.

3 Studies of the extent of hate speech in Norway

As demonstrated in the previous chapter, different studies produce divergent figures for the extent of hate speech, which is due partly to the fact that only a limited number of studies have operated with a precise definition of "hate speech" (see table 1). In this chapter, we will be seeking to arrive at a more reliable estimate of experiences of hate speech in Norway. We do so by analysing data from a survey conducted in June 2016 as part of the *Social Media in the Public Sphere* (SMIPS) project. In this survey, a large sample of the Norwegian population (n=5054) was asked if they had experienced being targeted by hate speech in social media. The question was accompanied by a definition of hate speech as "speech that is derogatory, threatening, harassing or stigmatising". Those who responded "Yes" were then given follow-up questions on which grounds the hate speech typically targeted. There was the option to select one or more of a total of 13 different grounds in addition to "Other" and "Don't know". The grounds included those protected by Section 185 of the Norwegian Penal Code (nationality, ethnicity, skin colour, religion, sexual orientation and ability), grounds which are typically included in more comprehensive definitions of hate speech (such as gender) (cf. Norwegian Equality and Anti-Discrimination Ombud 2015), and grounds not protected by law (such as occupation and education). It is important to emphasise that the data generated by the SMIPS survey is still being processed, and that more detailed analyses will be presented in a report to the Ministry of Justice and Public Security in connection with the project 'Prevention of online hate speech and hate crime'.[9]

In addition to the SMIPS survey, 'The Status of Freedom of Speech in Norway' project has in recent years conducted questionnaire-based surveys among the majority population, people of immigrant descent from Eastern Europe, Asia and Africa, and among journalists. Relevant findings from some of these surveys are cited in the previous chapter (Mitbøen & Steen-Johnsen 2016; Staksrud et al. 2014). This data has not, however, been put to full use, and in this chapter we present updated analyses. The questionnaires do not ask about "hate speech" specifically, but about "offensive or derogatory comments" and "threats". The aim is primarily to study relative differences between the

9 Due for publication in November/December 2016.

majority population on the one hand, and on the other, two groups we may assume are particularly exposed to hate speech and other distressing comments in social media, journalists and immigrants.

All of the surveys were conducted online by TNS Gallup among persons aged 15+. It is thus important to bear in mind that the findings are valid only for internet users in this age-group. The sample from the majority population was recruited from The Gallup Panel, which consists of randomly selected individuals who have agreed to complete questionnaires on a regular basis. The sample consisting of individuals of immigrant descent from Eastern Europe, Asia and Africa were recruited from the Norwegian National Registry (which in Norway forms the basis for the tax register, the electoral register and population statistics) among immigrants and their descendants resident in Norway for a minimum of five years. The sample of journalists was recruited from the member lists of the Norwegian Union of Journalists and the Association of Norwegian Editors.

As mentioned in Chapter 2, it is important to emphasise that questionnaire-based surveys on experiences of being the target of hate speech are *subjective* self-reports. We are thus not able to determine if the self-reported experiences would be defined as "hate speech" in the legal sense.

3.1 General population

We will start by looking at the extent of experiences of hate speech in the population as a whole. Table 3 indicates how many respondents in the SMIPS survey reported having been the targets of different forms of what they themselves perceived as hate speech, by gender. In addition to showing each individual ground, the table also sums up the proportion of the Norwegian population who has experienced hate speech directed at one of the grounds covered by Section 185 of the Norwegian Penal Code and an extended definition which also includes gender, appearance and personality.

Table 3. Grounds targeted by hate speech, as part of the population, by gender and in total. Percentage

	Men	Women	Total
Content of argument	**4.0**	**1.8**	2.9
Political stance	**3.9**	**1.6**	2.7
Personal attributes/personality	3.0	2.4	2.7
Appearance	1.2	1.0	1.1
Gender	0.5	1.6	1.1
Occupation	0.9	0.5	0.7
Nationality	**0.8**	**0.4**	0.6
Religion	**0.8**	**0.4**	0.6
Education	0.6	0.3	0.5
Disability/Ability	0.5	0.3	0.4
Skin colour	**0.6**	**0.2**	0.4
Sexual orientation	0.4	0.3	0.3
Ethnicity	**0.5**	**0.0**	0.2
Other	0.4	0.7	0.6
Don't know	0.4	0.6	0.5
Total	**8.4**	**5.9**	7.2
Grounds covered by Section 185	**2.7**	**1.2**	1.9
Section 185 + gender, appearance and personality	4.9	3.9	4.4
n (unweighted)	2611	2443	5054

Source: SMIPS 2016.
NOTE: Gender differences are highlighted where they are statistically significant ($p < 0.05$). Wording of questions: "Have you personally been the target of hate speech via social media? 'Hate speech' means speech that is derogatory, threatening, harassing or stigmatising." Respondents had the option of selecting multiple grounds. Section 185 provides protection against hate speech on the grounds of nationality, ethnicity, skin colour, religion, sexual orientation and ability. Weighted by gender, age and education.

A total of around 7 per cent of the respondents in the survey stated that they had experienced being the target of what they themselves perceived as hate speech. This response was given by more men (8.4 per cent) than women (5.9 per cent).

The table also shows that these instances of hate speech were typically directed at grounds not covered by the Norwegian Penal Code, and instead at the substance of their opinions, political stance and personal attributes/personality. Men reported more frequently than women that they had experienced hate speech targeting their political stance and the content of their argument, while

women responded more frequently than men that they had experienced hate speech targeting gender.

Considering only experiences of hate speech targeting nationality, ethnicity, skin colour, religion, sexual orientation and ability (Grounds covered by Section 185), the table shows that a total of 2 per cent of the respondents had experienced this. This is around the same level of hate crime more generally, as reported in the Norwegian Police population survey (Norwegian Directorate of Police 2016). More men than women reported on hate speech targeting the grounds covered by the Norwegian Penal Code.

If we expand the definition of hate speech to include gender, personality and appearance, the gender disparity disappears (is not statistically significant), and the overall scale of hate speech increases to just over 4 per cent. In other words, a large proportion of what people perceive as hate speech, falls outside of the definition in the Norwegian Penal Code – but also outside of more comprehensive definitions of the term.

When it comes to variations in sub-groups other than gender, the survey found that young people more frequently than older people have experienced being the target of hate speech, which is consistent with the studies discussed in the previous chapter (for example, Pew Research Centre 2014). There is also an apparent link between online behaviour and experiences of hate speech. People who regularly share their opinions and points of view on the internet and in social media especially are far more exposed than others to hate speech.

3.2 Individuals of immigrant descent from Eastern Europe, Asia and Africa

In 2013, the *Status of Freedom of Speech in Norway* project conducted surveys that permit comparison of the majority population with people of immigrant descent from Eastern Europe, Asia and Africa, in the age-group 16–50. The questionnaires in these surveys did not ask respondents directly about "hate speech", but about their experiences of receiving offensive or derogatory comments. Like the SMIPS survey, all respondents who answered "Yes" were asked a follow-up question on what the comments were typically directed at.[10] The respondents were then asked if they had also experienced receiving actual threats. Table 4 sums up the main findings of these surveys,

10 The list of grounds was shorter than in the SMIPS survey: disability, occupation and education were not included.

in which we distinguish between offensive or derogatory comments targeting different grounds.

Table 4. Experienced receiving offensive or derogatory comments and threats after participating in a discussion. The majority population and individuals of immigrant descent from Eastern Europe, Asia and Africa. Percentage

	Immigrants up to age 50 (2013)	Majority up to age 50 (2013)
Grounds covered by Section 185	**11.4**	**3.2**
Section 185 + gender, appearance and personality	12.7	10.2
Total offensive/derogatory comments	18.5	20.3
Actual threats	4.3	3.8
n (unweighted)	395	751

Source: Status of Freedom of Speech in Norway, 2013.
NOTE: Differences between the majority population and immigrants are highlighted where they are statistically significant ($p < 0.05$). Wording of questions: "After having participated in a discussion and stating your opinion publicly, have you ever received offensive or derogatory comments?" "Have you received actual threats against you personally?" Section 185 provides legal protection against hate speech directed at nationality, ethnicity, skin colour, religion and sexual orientation. Weighted by gender, age and education.

Table 4 shows that the overall extent of offensive or derogatory comments and threats was relatively similar for the two samples (the differences were not statistically significant). In both the majority sample and the immigrant sample, around 20 per cent responded that they had experienced receiving offensive or derogatory comments after participating in a discussion, while around 4 per cent responded that they had received actual threats.

However, if we restrict the offensive or derogatory comments to those targeting nationality, ethnicity, skin colour, religion and sexual orientation, we find a marked and significant difference between the samples. Not surprisingly, far more respondents of immigrant descent than respondents not of immigrant descent stated that they had experienced being the target of such comments. This is also consistent with surveys reviewed in the previous chapter (for example, Hatebase 2016).

However, if we extend the definition to include gender, appearance and personality, the difference disappears. A substantial proportion of the majority population has experienced receiving comments targeting these three grounds.

There are still rather more in the immigrant cohort than in the majority cohort who responded that they had personally been targeted by offensive or derogatory comments, but the difference is no longer statistically significant. In other words, the protected grounds are those that set the immigrants apart from the majority population.

3.3 Journalists and editors

In 2015, the *Status of Freedom of Speech in Norway* project conducted surveys that permit comparison of the majority population with members of the Norwegian Union of Journalists (Norsk Journalistlag – NJ) and the Association of Norwegian Editors (Norsk Redaktørforening – RF). To make the tables comparable, Table 5 shows only the proportion of the majority sample who responded that they had commented in at least one medium (including social media).[11]

Table 5. Experienced receiving offensive or derogatory comments and threats after participating in a discussion. Majority population and journalists. Percentage

	Members of NJ/RF (2015)	Majority – have stated an opinion (2015)
Grounds covered by Section 185	6.2	6.7
Section 185 + gender, appearance and personality	**32.6**	**16.2**
Total offensive/derogatory comments	**55.8**	**24.2**
Actual threats	**15.0**	**4.1**
n (unweighted)	1164	808

Source: Status of Freedom of Speech in Norway, 2015.
NOTE: Differences between the majority population and journalists are highlighted where they are statistically significant ($p < 0.05$). Wording of questions: "After having participated in a discussion and stating your opinion publicly, have you ever received any of the following? Offensive or derogatory comments – Threats". Section 185 provides legal protection against hate speech directed at nationality, ethnicity, skin colour, religion, sexual orientation and disability. Majority cohort weighted by gender, age and education.

Table 5 shows that more than half of Norwegian journalists (members of NJ and RF) have experienced being the target of offensive or derogatory comments. This is more than twice as many as in the majority sample, and indicates that

11 In these surveys, disability was included in the list of grounds. Occupation and education were not included.

this occupational group, not surprisingly, is at high risk of being exposed to offensive comments (cf. Hagen 2015).

If we restrict the comments to only those targeting nationality, ethnicity, skin colour, religion, sexual orientation and ability, the difference disappears. Compared with the majority population (who had expressed themselves in at least one medium), these figures thus suggest that journalists were not more subjected to offensive comments that would potentially be defined as illegal hate speech under Section 185.

However, journalists are exposed far more frequently than the general population to offensive comments directed at other personal attributes such as gender, appearance and personality. If we include these three attributes, the extent of offensive or derogatory comments is almost doubled in the journalist cohort compared to the majority cohort. A third of the journalists responded that they had experienced receiving comments targeting these attributes after participating in a discussion.

Finally, the table also shows that journalists are far more subjected to actual threats than the general population. 15 per cent of journalists responded that they had received threats, compared to 4 per cent of the majority population who had expressed themselves in at least one medium. Actual threats must be regarded as far more serious than offensive or derogatory comments, and in sum, the figures thus show that journalists as an occupational group are at risk.

3.4 Summary of research findings

In sum, the Norwegian surveys suggest that around 7 per cent of the population has experienced receiving what they perceive as hate speech. If we restrict the definition to apply only to speech targeting the grounds protected by Section 185 of the Norwegian Penal Code, the data suggests that around 2 per cent have experienced this. Further, we have seen that individuals of non-Western immigrant descent essentially have as much experience of offensive or derogatory comments and threats as the general population has, but that the content of the comments is different. The immigrants who participated in the survey had far more frequently experienced receiving comments targeting one of the protected grounds. Finally, we have seen that a very high proportion of journalists report experiencing offensive or derogatory comments and threats. However, the journalists are not more subjected to comments targeting the protected grounds than the general population. Instead, they receive a very large volume of

comments targeting aspects such as gender, appearance and personality, in addition to the content of their argument and political stance.

It is important to emphasise that the surveys reviewed in this chapter only study individual experiences of being targeted by different types of speech directly. As stated earlier in the report, Section 185 of the Norwegian Penal Code also provides protection against speech conveyed publicly without it targeting a specific individual. The reason given for this is partly that the fact of witnessing hate speech can also have adverse consequences. In the systematic review in Chapter 2, we saw that the proportion who witnessed hate speech or offensive comments is far higher than the proportion who were at the receiving end of such speech or comments (Hawdon et al. 2015; Pew Research Center 2014).

4 Perpetrators of online hate

So far, we have concentrated on the extent of hate speech and its targets. In this chapter we turn our attention to what characterises the perpetrators of online hate. The number of quantitative studies specifically on the perpetrators of online hate is limited. However, there are a number of studies that indirectly characterise the perpetrators by asking their victims to describe them. In addition, several studies have in various ways studied the characteristics of perpetrators and their motives. Yet others have examined this trend in online culture in more general terms.

There is much to indicate that a strong feeling of hate is not in itself a prime motive for hate speech, and factors such as thrill-seeking and an internet culture of defamatory language are more likely the key to understanding the phenomenon. However, the victims of attacks from those who may simply be 'bored' are not random. Social, structural and cultural trends influence our perception of what and who is "different". Perpetrators base their actions on prejudices, stereotypes and assumptions about differences between groups (Chakraborti & Garland 2015), and hate speech should thus be understood as an expression of this.

4.1 Individual characteristics of the perpetrators of hate speech

The research literature has examined what characterises the perpetrators of hate speech, harassing messages and hate crime. Here we present some of the main research findings.

- *Relationship with the victim*: The perpetrators of online hate speech are not necessarily strangers to their victims. In the past, hate crime was seen as 'stranger danger', where the victim is a random member of a minority group. However, recent research reveals that there is often some form of relationship between the victim and perpetrator, especially when it comes to less serious forms of hate crime such as harassment (Chakraborti & Garland 2015). Half of those subjected to online harassment reported that the perpetrator was someone they knew (Pew Research Center 2014: 26).

- *Gender*: Men are more frequently behind online hate than women, and men exhibit greater tolerance of online hate and sexual harassment than women (Ask et al. 2016; Hagen 2015).
- *Social class:* When it comes to social class, the picture of the perpetrators of hate speech is more complex. While Potok (2016) contends that "working class" people are over-represented among the perpetrators of online hate speech, other studies show that people with relative high social status also commit hate crimes (Chakraborti & Garland 2015; Perry 2001). There is thus reason to believe that the social class of those behind hate speech is mixed, but also that this varies between societies and cultures.
- *Age:* No consistent picture emerges for age. A study of the profile of 169 people convicted of hate crimes in Boston, USA found that young people are over-represented (McDevitt, Levin& Bennet 2002). These were, however, serious cases, in which the perpetrators had been convicted. It would appear that the age profile has more spread when it comes to less serious cases. A study of racist harassment and violence, for example, found that all age groups are involved (Sibbitt 1997 in Hall 2005. 87). As we will return to below, it would also appear that verbal abuse targeting gender and sexual orientation is mainstream in youth culture generally (Helseth 2007).
- *Personality:* A study of the link between the style of online commentary and personality show that trolling (anonymous destructive and offensive behaviour) is linked to sadistic and psychopathic personality traits (Buckels, Trapnell & Paulhus 2014).
- *Individuals rather than organised groups:* American research in particular has focused on organised hate, that is, on online hate groups and hate sites. However, the prevalence of online hate groups has diminished over the last five years, and the primary sources of online hate material appear now to be individuals (Potok 2016). Norway does not have the same history of hate groups as in the US.

4.2 The motives for hate speech and hate crime

The motivation for perpetrating hate speech or hate crime varies. Several studies have produced typologies of the perpetrators and their motivation (for example Erjavec & Kovačič 2012; McDevitt et al. 2002), and here we will present a summary of their findings.

- *Thrill-seeking and boredom:* For some, the main motivation for posting hate speech or insults or perpetrating hate crimes is sheer thrill-seeking, attention-seeking and personal amusement (Buckels et al. 2014; Erjavec & Kovačič

2012; Shachaf & Hara 2010). A number of perpetrators see it as a ‘game’, which leads Jane (2014) to contend that much of what appears to be hate speech, should instead be seen as “boredom speech”.

- *Education and information:* People who post extremist and hateful comments do not necessarily see them as offensive. The perpetrators may regard their content as informative, and the main motive for posting it online is to spread awareness of their group or ideology, to defend their group against criticism or recruit others to “their” cause (Hawdon et al. 2015; McNamee, Peterson & Peña 2010).
- *Reinforce group identities:* The internet gives people with extremist views an outlet for their points of view and frustrations. Extremist groups use the internet to connect with people with similar attitudes and in that way achieve a sense of shared identity and common goals (Gerstenfeld, Grant & Chiang 2003). For some, the motivation for disseminating hate speech is to defend their group interests and attack “the enemy”. They regard their online “war” as an extension of their lobby in politics and society generally (Erjavec & Kovačič 2012).
- *Hate:* Some have a well-defined ideological motivation and are driven by sheer hate against a group. A review of hate crime offenders, however, found that it was extremely rare for the crime motive to be sheer hate (McDevitt et al. 2002). This means that hate speech should not be understood as being motivated purely by strong feelings of hate.

4.3 Cyberculture

On the one hand, the internet represents a democratisation of access to the public arena because it is accessible to “everyone” and thus increases the scope for exercising freedom of speech. On the other hand, when seeking to account for offensive online behaviour, it is customary to rely on certain characteristics of the internet as an arena: the internet offers a sense of anonymity which makes people feel that they can make extreme statements without detection and without accountability (Lapidot-Lefler & Barak, 2012 in Ask et al. 2016; Foxman & Wolf 2013; Gagliardone, Gal, Alves & Martinez 2015). There is little empirical evidence that anonymity inherently spurs misconduct or anti-normative behaviour online (Douglas 2007; Santana 2014), but there are studies indicating that debates among anonymous participants are less “civilised” than debates using real names (Santana 2014).

Several researchers make the point that some segments of cyberculture have developed a culture of high tolerance of negative and offensive statements. Jane (2014) for example argues that hyper-aggressive and nasty language has become the modus operandi and almost expected in some environments in expressing dissension, testing and staking out limits in online communities, competing, relieving boredom, attention-seeking and/or sheer thrill-seeking (Jane 2014: 542).

Norwegian research on online gaming points to the same: gaming culture has a high level of acceptance of referring to minority groups in derogatory terms, but the gamers do not make much of it. The terms of abuse are "just" a way of communicating (Ask et al. 2016: 11). At the same time, verbal abuse targeting gender and sexual orientation would also appear to be a trend in youth culture generally (Helseth 2007).

4.4 Larger social and cultural structures

Within research on hate crime, hate crime is commonly understood as an expression of social and collective frustrations. When minorities are perceived as a threat to achieving personal aims, for example, because of an impression that they "take jobs from ordinary people", hate crime may be the response (see for example Glick 2005; Hall 2005).

Perry (2001) contends that hate crime should be understood as a way of "*doing difference"*. A number of hierarchical power structures in society are based on preconceptions of differences between groups, with one group representing "the norm" on top, and those seen as "different" assigned to subordinate positions. It is typically when people cross or threaten boundaries and "forget their place" that hate crime is perpetrated in response to the threat. Hate crime or hate speech is a way of staking out boundaries between groups. It serves to remind those who are "different" where they belong. Hate speech and hate crime can thus serve as a tool in attempting to maintain or reaffirm a (perceived) dominance. Perpetrators "re-create their own masculinity, or whiteness, for example, while punishing the victims for their deviant identity performance" (Perry 2001: 55).

This perspective on hate speech and hate crime diverts attention away from the individual perpetrator to the social context and power relations that otherwise exist in society. The hate rhetoric in public debate is consequently regarded as an expression of broader cultural perceptions of differences between groups, prejudices and stereotypes.

An American study may serve to elucidate the question of how hate rhetoric in public debate is influenced by wider debates in society such as on immigration, religion, political conflicts and gender equality topics. A team of researchers introduced various types of scenarios in white supremacist chat rooms. The scenarios represented different levels of "threats": low (competition for jobs); moderate (immigration); high (interracial marriage). The researchers also manipulated the scenarios so that the threats targeted the personal, community or national level. The study revealed that the respondents only resorted to inciting violence in response to scenarios perceived as a high and personal threat (for example, first-hand experience of interracial marriage) (Glaser et al. 2002 in McNamee et al. 2010: 259).

This suggests that certain scenarios may be perceived as threats and that these scenarios have the potential to trigger hate speech and incitement to violence. Any strong emphasis that minority groups pose a threat (for example in relation to immigration or gender equality) may thus potentially increase the scale of hate speech.

Another aspect of this topic is addressed in Report 2 (Eggebø & Stubberud 2016): that vitriolic and hate-ridden rhetoric in public debate can have the effect of shifting the boundaries of what are regarded as acceptable expressions and in that way pave the way for even more vitriolic hate speech.

5 Research status and needs

The purpose of this report has been to gather research-based knowledge concerning:

- the extent of online hate speech
- which groups in society are particularly subjected to online hate speech
- who produces hate speech, and what motivates them

The report demonstrates that statistics on officially reported instances of alleged criminal hate speech are extremely limited and not adequate for determining the extent of hate speech. Various questionnaire-based surveys serve to fill out the picture, but one of the problems of this type of research is that very few studies have investigated the extent of hate speech directly. Different studies have applied different definitions of hate speech and other offensive statements and hence arrived at different conclusions. Many of the studies have primarily addressed online hate more generally, without reference to the legal concept of "hate speech". This means that there are few studies that examine hate speech aimed at specific group identities.

Regardless of how the prevailing terms have been defined in the empirical studies, our systematic review shows that ethnicity and sexual orientation are highly prevalent grounds for hate speech, while less research has been conducted on online hate targeting disability.[12] Furthermore, the young are far more often subjected to offensive experiences online than are the elderly. Overall, gender differences are relatively insignificant when it comes to exposure, but women and men are subjected to different forms of online hate.

Analyses of a Norwegian survey conducted in June 2016 show that around two per cent of the population respond that they have been targets of hate speech in social media targeting the grounds protected by Section 185 of the Norwegian Penal Code. Disregarding which grounds that hate speech is actually targeting, our analyses show that around seven per cent of the Norwegian population reported being the targets of what they themselves *perceive* as hate speech. However, most of these instances of hate speech target grounds that are not covered by Section 185.

12 However, a report on hate speech targeting people with disabilities will be published in autumn 2016 (Olsen et al. 2016).

The analyses show that individuals of immigrant descent from Eastern Europe, Asia and Africa are more subjected than the majority population to offensive and derogatory comments targeting the grounds covered by Section 185, but that the majority population has more frequently been subjected to offensive comments on other grounds. Journalists are generally far more subjected than the general population to offensive or derogatory comments and threats. However, our analyses do not suggest that journalists are more at risk of comments targeting the grounds protected by Section 185; on the contrary, they receive a great many offensive comments targeting other grounds.

Our systematic review of research on the perpetrators of online hate speech has shown that these are typically men, and also that men have a greater tolerance of online hate than women do. Moreover, the perpetrators are also commonly motivated by factors other than a strong feeling of hate. Factors such as thrill-seeking and an internet culture of defamatory language are more likely key to understanding the underlying dynamics of hate speech. That said, the victims of hate speech are not random. The perpetrators base their views on prevailing prejudices, stereotypes and subjective assumptions about differences between groups in society, and hate speech is directed at those perceived as "different".

5.1 Research needs

Several actors have indicated a need for substantial research on hate speech and hate crime targeting minority groups in Norway. Among these are the authors of a systematic review of research on discrimination against Sami people, national minorities and immigrants in Norway (Midtbøen & Lidén 2015). Similarly, a systematic review of research on radicalisation and violent extremism concludes that there is an obvious need for more research on online hate rhetoric and threats, and on hate crime in general (Bjørgo & Gjelsvik 2015: 252; see also LDO (Norwegian Equality and Anti-Discrimination Ombud) 2015).

Based on the systematic review in this report, in addition to the review in Report 2 (Eggebø & Stubberud 2016), we have identified four primary research needs: 1) research on the extent and experiences of hate speech, including discriminatory and offensive statements; 2) research on the perpetrators of hate speech; 3) textual analysis of hate speech and discriminatory statements in the public sphere and 4) research on the consequences of hate speech.[13]

13 The description of these research needs is the same for Report 1 and Report 2.

The extent and experiences of hate speech, and discriminatory and offensive statements

There is a need for more empirical research on the extent of different forms of hate speech and discriminatory statements, and on the experiences of being targeted by such speech and statements among potentially vulnerable groups. Such research studies should cover the following issues:

- *Varying degrees of "hate"*: There is little consonance in terms of terminology in research on hate speech. One essential research need is therefore to produce a more detailed overview of different forms of "hate". "Hate speech" as a concept is difficult to measure accurately, and research is required to study the experiences of different forms of hate speech, discriminatory and other offensive statements, within a consistent survey design. Research on hate speech and hate crime should be analysed within the context of general research on discrimination. New studies should be designed so that they are able to chart a wide range of different experiences of discrimination, including hate speech and hate crime (see for example the set of indicators applied by Andersen, Buer, Olaniyan & Malterud 2016).

- *Protected grounds:* Hate speech targets different minority groups in society. Existing survey research is mainly based on representative samples of the general population, in which minority groups naturally make up a small proportion. In Norway, specific surveys have been conducted among persons with disabilities (Olsen et al. 2016) and individuals of immigrant descent from Eastern Europe, Asia and Africa (Midtbøen & Steen-Johnsen 2016; Staksrud et al. 2014). The latter survey had few respondents and a low response rate. There is therefore a need for better-designed, more comprehensive and more representative surveys among individuals of immigrant descent. Furthermore, there is a need to study other minority groups. There is very little research to shed light on hate speech – or discrimination in general – targeting Norway's indigenous people (Sami) and national minorities (Kven people, Forest Finns, indigenous travellers/Romani people, the Roma people in Norway and Norwegian Jews). There is a large body of research on LGBT persons (lesbian, gay, bisexual and transgender), but no Norwegian studies specifically investigating hate crime or hate speech targeting this group.

- *Other at-risk groups*: As discussed at various points in this systematic review, in many cases it is relevant to operate with a broader definition of hate speech than that described in the Norwegian Penal Code (see also LDO (Norwegian Equality and Anti-discrimination Ombud) 2015). There is hence a need to

extend research to experiences of hate speech and other discriminatory and offensive statements targeting other grounds than those afforded legal protection, such as gender, gender expression, gender identity and age. There is also a need to study hate speech within the context of the labour market. To date, the only studies carried out of the labour market have concerned journalists (Hagen 2015), but it will be of relevance to study other professions that are potentially at risk, such as teachers, street-level bureaucrats in the public sector and politicians.

- *Interaction between different group identities*: Research to date has largely described experiences of hate speech or other offensive statements one group at a time. Research on sexual harassment, however, shows that there may be an interaction between different group identities that increase an individual's vulnerability. Sexual harassment, for example, is closely linked to harassment on grounds of both gender and sexual orientation, racist harassment and harassment of people with disabilities (Ask et al. 2016; Buchanan & Fitzgerald 2008; Shaw, Chan & McMahon 2012). Studies of experiences of hate speech should therefore not focus on a single group identity at a time, but rather investigate the interaction between different group identities.

- *Differentiation of arenas*: Research on hate speech should not be restricted to the internet as arena. Empirical studies of the phenomenon should allow for differentiation between different arenas, for example the public sphere, schools, social media, at work etc. The different arenas in which hate speech is perpetrated may have different consequences for the individual or group targeted.

- *Differentiation of victims*: Existing research has mainly focused on the experiences reported by single individuals of being the target of hate speech and other offensive statements. As discussed in Report 3 (Wessel-Aas et al. 2016), one key provision in Section 185 of the Norwegian Penal Code is that hate speech does not necessarily have to directly target one person, but that hate speech conveyed publicly and more generally ("in the presence of others") is also a criminal offence. Future research should therefore also study the experiences of individuals who are victims of hate speech, and their experience of witnessing hate speech conveyed publicly.

Perpetrators of hate speech

- *Who are the perpetrators?* Little research has been done on who the perpetrators of hate speech are. Research on perpetrators is important, as it provides both a better understanding of hate speech as a phenomenon and a better understanding of the factors underlying hate speech. Such research will be of particular value for targeted measures to prevent hate speech. Producing more research on perpetrators poses methodological challenges, such as the fact that many perpetrators are unwilling to admit or acknowledge that they are conveying hate speech. Methods have been developed to study such sensitive issues when using questionnaires, for example, using "List Experiments". This is a method developed to measure controversial conduct and controversial attitudes without the respondents having to admit to such conduct and attitudes directly (see for example Mutz 2011: chapter 2).

- *What motivates the perpetrators?* In the interests of prevention, it is essential to gain a better understanding of the motivation for hate speech. Why do they say what they do, and how do they perceive their own statements? (see also Gagliardone et al. 2015: 57). Does hate speech correlate with negative attitudes and prejudices towards specific groups in the population? Have the perpetrators experienced bullying, either as bullies themselves or victims of bullying? Are the perpetrators of hate speech more likely than others to commit other more serious forms of hate crime? And in extension of this; is there a link between hate speech and other forms of hate crime?

Textual analyses

- *The contents of hate speech*: Research on hate speech is predominately questionnaire-based. Only very few studies have actually examined the opinions expressed in hate speech. An analysis of such online content will allow us to study actual hate speech – in order to gain knowledge of its content and perpetrators, estimate the extent of hate speech and analyse changes over time. An approach of this type can also advance our understanding of the relationship between public discourse and hate rhetoric. Do hate speech and other offensive statements in the public domain limit or increase the extent of such expressions in social and conventional media?

- *Speech culture*: One related field is the study of what has been termed "speech culture" within different (sub)public spheres. How does hate rhetoric vary according to different arenas, such as closed groups and open debate fora – in public discourse? Furthermore: How do people perceive the boundaries between different (sub)public spheres? Is there a higher tolerance

of hate speech and other offensive statements on certain arenas? What is the mutual impact of debate culture vis-à-vis different arenas?

Consequences of hate speech

- *Differentiated consequences*: Norwegian studies to date have in practice exclusively addressed a single type of consequence of experiencing hate speech or offensive statements, i.e. the likelihood that persons subjected to such statements will withdraw from public debate. In other words, the studies have focused on whether experiencing hate speech and other offensive statements have impacted the willingness to exercise freedom of speech. Hate speech can, as illustrated in Report 2 (Eggebø & Stubberud 2016), have a number of other consequences for individuals, groups and society at large, and there is a need for more research into the consequences at all three levels.

- *A changing public sphere*: The debate on freedom of speech comprises varying opinions on what constitutes the best protection against hate speech. On the one hand is the contention that the best protection against hate speech and other offensive statements is to allow them space in public discourse so that they can be criticised and countered ("the concept of the cleansing function of public debate"). On the other hand it is claimed that hate speech can have the effect of shifting the boundaries for what is perceived as legitimate expression in public. We have, however, little empirical evidence of how such mechanisms would work in practice. Would the population become 'hardened' by increased exposure to hate speech and other offensive statements, or would exposure serve to mobilise counterforces?

- *Prevention*: A systematic review of relevant literature should be conducted, to shed light on the breadth and impact of various preventive measures implemented in Norway and comparable countries. A review of prevention should include literature, methodological handbooks and project reports. The review should be comprehensive and not restricted to measures in the form of legislation. The systematic review should equally not be limited to hate speech, but examine measures against related phenomena such as discrimination and bullying. Examples of measures to be studied should include attitudinal campaigns, measures offered to help victims, measures targeting perpetrators and efforts aimed at specific local communities.

Bibliography

Aftenposten (2016, 8.februar 2016). – De skrev at jeg var stygg og ikke fortjente å leve, *Aftenposten*. Hentet fra http://www.aftenposten.no/nyheter/iriks/--De-skrev-at-jeg-var-stygg-og-ikke-fortjente-a-leve-8334905.html

Andersen, N., Buer, L., Olaniyan, O.S. & Malterud, K. (2016). *Kartlegging av levekår blant lesbiske, homofile, bifile og transpersoner. Et indikatorsett for ti års oppfølging.* Bergen.

Anderssen, N. & Malterud, K. (red.) (2013). *Seksuell orientering og levekår*. Bergen: Uni Helse/Uni Research.

Ask, K., Svendsen, S.H.B. & Karlstrøm, H. (2016). Når jentene må inn i skapet: Seksuell trakassering og kjønnsfrihet i online dataspill. *Norsk medietidsskrift, 22*(01).

Bjørgo, T. & Gjelsvik, I.M. (2015). *Forskning på forebygging av radikalisering og voldelig ekstremisme: En kunnskapsstatus.*Rapport nr. 2015:2. Oslo: Politihøyskolen.

Bleich, E. (2011). *The freedom to be racist?: how the United States and Europe struggle to preserve freedom and combat racism*. Oxford: Oxford University Press.

Buchanan, N.T. & Fitzgerald, L.F. (2008). Effects of racial and sexual harassment on work and the psychological well-being of African American women. *Journal of occupational health psychology, 13*(2), 137.

Buckels, E.E., Trapnell, P.D. & Paulhus, D.L. (2014). Trolls just want to have fun. *Personality and Individual Differences, 67*, 97–102.

Chakraborti, N. & Garland, J. (2015). *Hate crime: impact, causes and responses* (2. utg.). Los Angeles: Sage.

COWI. (2015). *Kortlægning af hadforbrydelser i Danmark. En undersøgelse af befolkningens oplevede hadforbrydelser*. Kongens Lyngby.

Digranes, E. (2016). *Disabled justice: A qualitative case study on barriers to registration of disability hate crime within the Norwegian criminal justice system.* Masteroppgave. Oslo: Universitetet i Oslo.

Djuve, A.B., Gulløy, E., Kavli, H.C. & Berglund, F. (2009). *Datafangst når minoritetsbefolkningen er målgruppe*. Fafo-rapport 2009:24. Oslo.

Douglas, K.M. (2007). Psychology, discrimination and hate groups online. *The Oxford handbook of internet psychology*, Oxford: Oxford University Press.155–164.

Eggebø, H. & Stubberud, E. (2016). *Hate speech, report 2: Research on hate and discrimination*. Oslo: Institute for Social Research.

Erjavec, K. & Kovačič, M.P. (2012). “You Don’t Understand, This is a New War!” Analysis of Hate Speech in News Web Sites’ Comments. *Mass Communication and Society, 15*(6), 899–920.

Foxman, A.H. & Wolf, C. (2013). *Viral hate: Containing its spread on the Internet*. New York: Macmillan.

Gagliardone, I., Gal, D., Alves, T. & Martinez, G. (2015). *Countering Online Hate Speech*. Paris: UNESCO

Gerstenfeld, P.B., Grant, D.R. & Chiang, C.-P. (2003). Hate Online: A Content Analysis of Extremist Internet Sites. *Analyses of Social Issues and Public Policy, 3*(1), 29–44.

Ghersetti, M. (2015). Sociala medier till nytta och nöje. I A. Bergström, B. Johansson, H. Oscarsson & M. Oskarson (red.), *Fragment*. Gøteborg: SOM-institutet.

Glick, P. (2005). Choice of Scapegoats. I J.F. Dovidio, P. Glick & L.A. Rudman (red.), *On the nature of prejudice: fifty years after Allport* (s. 244–261). Malden, Mass: Blackwell.

Hagen, A.L. (2015). *Meningers mot: netthat og ytringsfrihet i Norge*. Oslo: Cappelen Damm Akademisk.

Hall, N. (2005). *Hate crime*. Cullompton: Willan Publishing.

Hall, N. (2013). *Hate crime* (2. utg.). London: Routledge.

Hatebase. (2016). Most common hate speech. Hentet 15.08.2016, fra https://www.hatebase.org/popular

Hawdon, J., Oksanen, A. & Räsänen, P. (2015). Online Extremism and Online Hate: Exposure among Adolescents and Young Adults in Four Nations. *Nordicom Information: Medie- och kommunikationsforskning i Norden, 37*(3–4), 29–37.

Helseth, H. (2007). *Kunnskapsstatus om kjønnsrelatert mobbing blant barn og unge*. Oslo: Høgskolen i Nesna.

Hoffmann, C., Kopperud, Ø. & Moe, V. (2012). *Antisemittisme i Norge?: den norske befolkningens holdninger til jøder og andre minoriteter*. Oslo.

Jane, E.A. (2014). “Your a Ugly, Whorish, Slut”. *Feminist Media Studies, 14*(4), 531–546.

Livingstone, S., Haddon, L., Görzig, A. & Ólafsson, K. (2011). *Risks and safety on the internet: the perspective of European children: full findings and policy implications from the EU Kids Online survey of 9–16 year olds and their parents in 25 countries*nr. Deliverable D4. London.

McDevitt, J., Levin, J. & Bennett, S. (2002). Hate crime offenders: An expanded typology. *Journal of Social Issues, 58*(2), 303–317.

McNamee, L.G., Peterson, B.L. & Peña, J. (2010). A call to educate, participate, invoke and indict: Understanding the communication of online hate groups. *Communication Monographs, 77*(2), 257–280.

Midtbøen, A.H. & Lidén, H. (2015). *Diskriminering av samer, nasjonale minoriteter og innvandrere i Norge: En kunnskapsgjennomgang*. Oslo: Institutt for samfunnsforskning.

Midtbøen, A.H. & Steen-Johnsen, K. (2016). Ytringsfrihetens grenser i det flerkulturelle Norge. *Nytt norsk tidsskrift, 33*(1–2), 21–33.

Mutz, D.C. (2011). *Population-based survey experiments*. Princeton & Oxford: Princeton University Press.

Mølster, R. (2015). Hat og trusler mot journalister i Skandinavia. *Nordicom Information: Medie- och kommunikationsforskning i Norden, 37*(3–4), 57–62.

Nilsen, A.B. (2014). *Hatprat*. Oslo: Cappelen Damm Akademisk.

Nilsson, M.L. (2015). Hot och hat mot svenska journalister. *Nordicom Information: Medie- och kommunikationsforskning i Norden, 37*(3–4), 51–56.

Norwegian Directorate of Police (2012). *Politiets nasjonale innbyggerundersøkelse 2012*.

Norwegian Directorate of Police (2015). *Anmeldelser med hatmotiv, 2010–2014*.

Norwegian Directorate of Police (2016). *Politiets innbyggerundersøkelse*.

Norwegian Equality and Anti-Discrimination Ombud [Likestilings- og diskrimineringsombudet] (2015). *Hate speech and Hate*. Oslo: Norwegian Equality and Anti-Discrimination Ombud.

Norwegian Media Authority (2016). *Barn & medier 2016: 9–16-åringers bruk og opplevelser av medier*. Fredrikstad: Medietilsynet.

NOVA (2015). *Ungdata: Nasjonale resultater 2014*. NOVA rapport 7/15. Oslo.

Olsen, T., Vedeler, J., Eriksen, J. & Elvegård, K. (2016). *Hatytringer. Resultater fra en studie av funksjonshemmedes erfaringer*. Rapport nr. 6/2016. Bodø: Norlandsforskning.

Oslo Police District (2013). *Hatkriminalitet: En drøfting av temaet, og gjennomgang av anmeldelser i Oslo 2012*.

Oslo Police District (2016). *Hatkriminalitet: Anmeldt hatkriminalitet Oslo politidistrikt 2015*.

Perry, B. (2001). *In the name of hate: understanding hate crimes*. New York: Routledge.

Pew Research Center (2014). *Online Harassment*.

Potok, M. (2016). *The year in hate and extremism 2016*.

Ramm, J. (2010). *På like vilkår? Helse og levekår blant personer med nedsatt funksjonsevne*. Oslo/ Kongsvinger: Statistisk sentralbyrå.

Roland, E. & Auestad, G. (2009). *Seksuell orientering og mobbing*. Stavanger: Senter for atferdsforskning.

Santana, A.D. (2014). Virtuous or Vitriolic. *Journalism Practice, 8*(1), 18–33.

Shachaf, P. & Hara, N. (2010). Beyond vandalism: Wikipedia trolls. *Journal of Information Science, 36*(3), 357–370.

Shaw, L.R., Chan, F. & McMahon, B.T. (2012). Intersectionality and Disability Harassment: The Interactive Effects of Disability, Race, Age, and Gender. *Rehabilitation Counseling Bulletin, 55*(2), 82–91.

Staksrud, E., Steen-Johnsen, K., Enjolras, B., Gustafsson, M.H., Ihlebæk, K.A., Midtbøen, A.H. mfl. (2014). *Status for ytringsfriheten i Norge. Resultater fra befolkningsundersøkelsen 2014*. Oslo: Fritt Ord, ISF, IMK, Fafo.

Sunde, I.M. (2013). *Forebygging av radikalisering og voldelig ekstremisme på internett*. 2013:1. Oslo: Politihøyskolen.

Wessel-Aas, J., Fladmoe, A. & Nadim, M. (2016). *Hate speech, report 3: The boundary between freedom of speech and criminal law protection against hate speech*. Oslo: Institute for Social Research.

Institutt for
samfunnsforskning

Institute for
Social Research

Munthes gate 31
PO Box 3233 Elisenberg
NO-0208 Oslo, Norway
Tel +47 23 08 61 00
samfunnsforskning.no

ISBN (online): 978-82-7763-539-2
ISSN (online): 1891-4314