Abuse and anonymity

2022

A report for the Department for Digital, Culture, Media & Sport

REVEALING REALITY

Department for
Digital, Culture
Media & Sport

About Revealing Reality

Revealing Reality is an independent social research agency, working with regulators, government and charities to provide insight into people's online behaviours and experiences.

Studying how the digital world is shaping people's behaviours is something we do every day. We regularly conduct detailed qualitative behavioural research, observing how people really use digital products, services and technology. This includes exploring how digital design shapes behaviour – across technology, gambling, financial products, the health service, and more.

Visit www.revealingreality.co.uk to find out more about our work or to get in touch.

Contents

Executive summary

In autumn 2021, the Department of Digital, Culture, Media and Sport (DCMS) commissioned Revealing Reality to carry out research to investigate any links between anonymity and online abuse.

This followed the publication of the draft Online Safety Bill in May 2021, which included measures to ensure services in scope addressed anonymous abuse on their platforms, after the government response to the Online Harms White Paper in 2020 highlighted the impact anonymous abuse could have on members of the public or high-profile figures.

At the point at which this research was commissioned, evidence was mixed as to whether being anonymous online led to more people being more abusive.

This research was intended to feed into the ongoing development of the Online Safety Bill through its provision of:

1. A rapid review of the existing literature and evidence on the potential relationship between online abuse and online anonymity.
2. Primary, qualitative research of a kind that had not previously been carried out, to provide additional, new evidence of that potential relationship.

While this research project was underway, the development of the Online Safety Bill continued and it was introduced to Parliament in March 2022. This iteration of the Bill included provisions that had not been included in previous drafts.

These included measures to help adults control the content they saw from, and the interaction they have with, other users. The government has consistently recognised the importance of anonymity online, saying that removing anonymity would have negative consequences for some people, for example, who used anonymity online for their personal safety or to explore their sexuality.

This decision chimes with many of the findings in this report, namely that:

- online anonymity benefits some users
- online anonymity is not in and of itself the main or only enabler of online abuse

As a result of these findings, it would appear that banning anonymity online would have negative consequences and would not solve the problem of online abuse.

This report provides an overview of the evidence and thinking that informed that decision.

It also provides a wider understanding of the ways in which online abuse, including anonymous abuse, can come about and be perpetuated. This will help inform the ongoing need for online companies, users and policy-makers to develop and apply mechanisms and considerations that could help reduce or potentially prevent future abuse online.

We hope that the research findings and conclusions will also help stakeholders and wider audiences understand the limitations of such measures, and consider what further research would help inform this all-reaching area of policy and experience.

Previous research was not able to conclude that removing anonymity would reduce the impact of online abuse overall

The relationship between online abuse and anonymity was contested in previous research[1].

Several large-scale studies suggested allowing online anonymity increased the likelihood of abusive behaviours. However, much of the previous research was limited in its methodology or scope, and used differing definitions of abuse, limited sampling criteria, and/or did not account for the severity of abuse or other contextual information.

---

[1] See full evidence review in Annex for details of research that explores this relationship

Despite some studies suggesting anonymity did lead to more or worse abuse, removing anonymity was rarely suggested as the best means to reduce it. Reasons for this included:

- abuse still occurs from non-anonymous users
- sometimes these users *prefer* having identifiable accounts for being abusive
- anonymity was highly valued by many users for reasons that were nothing to do with abuse and who were not abusive

Previous work lacked research on the impact of anonymous abuse compared with abuse from identifiable perpetrators. Understanding the impact of anonymous abuse was essential to calculate how important it was to deploy resources to reduce it, compared with other measures.

Some studies explored the culture on different platforms, or how social norm theory related to anonymity, referencing experiments or platforms' attempts to change behaviour through, for example, platform design or interventions by moderators.

In countries that have taken steps to reduce anonymity online, there is so far little data on how effective those measures have been. In South Korea, for example, there was little evidence of impact, and the government reversed the policy.

Many users felt some approaches to reducing anonymity online were safe and acceptable, but there was limited evidence that it would have a positive effect overall in reducing abuse, including being certain that any reduction in anonymous abuse would not result in an increase in abuse of other kinds.

Reviewing this body of previous work[2], Revealing Reality concluded more research was needed that:

- specifically explored the relationship between abuse and anonymity
- used clear and consistent definitions of abuse and anonymity online
- explored the motivations of users posting abuse
- sought to understand the impact that online abuse, including anonymous abuse, has on people

This qualitative research provides new evidence of people's real-life experiences of abuse and anonymity online

To address gaps in previous work, and provide evidence as to whether being anonymous online led to more people being more abusive, Revealing Reality interviewed 12 relevant subject-matter experts and 23 users of social media who were victims of abuse, perpetrators of abuse, valued anonymity online or, in several cases, fell into more than one of these categories.

The expert interviews, which took place online, were with:

- 4 representatives of social media platforms about their policies and activity to protect people from abuse and / or respond to it
- 4 representatives of organisations supporting victims of online abuse exploring real-life examples and trends
- 5 representatives of organisations or individuals who have a valuable perspective on anonymity online, for example policy professionals

The in-depth interviews with users lasted up to two-hours and were conducted face to face, via video call, and on written chat, when respondents wanted to remain completely anonymous.

As a small, qualitative study, this research cannot be used to quantify experiences or definitively prove or disprove a link between anonymity and abuse. However, this qualitative research reveals people's real-life experiences and behaviours, adding nuance and fresh insight to this complex area.

During the interviews, the researchers explored with respondents:

- What different types of abuse and anonymity looked like online
- Whether and how experiences of abuse were linked to anonymity
- How anonymous and non-anonymous abuse affected victims

---

[2] See full rapid literature review on p.54

- How factors other than anonymity shaped abusive behaviour
- How some people were benefiting from anonymity.

Findings from this research

Data and analysis from the interviews led to several key findings.

Isolating the role anonymity plays in abusive behaviour is impossible in practice

There appears to be a relationship between anonymity and abuse in some cases but not others. Even where it looks as though there is a link, isolating the role that anonymity plays in facilitating or magnifying the abuse is practically impossible. In addition, there are usually other factors at play, as detailed below. It is difficult to disentangle which of these has the greatest impact, and the impact is likely to change depending on the circumstances.

Even with a much larger sample than this qualitative research allowed for, it would be incredibly difficult to determine a causal link between anonymity and abuse given the differing influence in each case of platform design, the behaviour of other users, and the sensibilities of those involved.

There are different types of abusive behaviours and content, different degrees of anonymity and a disposability of social media accounts, which can make it easier for perpetrators to post abuse

There are different types of abusive behaviours and content online, and the impact of these differs. Among the experts and users we spoke to, these behaviours tended to be one of:

- Reactive, angry or one-off
- Sustained or obsessive – closer to harassment
- Driven by ideology or an agenda
- Intended to discredit or 'call out' individuals

Anonymity cannot be understood in binary terms. There are different degrees of anonymity, which can have different implications for users and should not be conflated:

- Being completely identifiable
- Being anonymous to other users but identifiable to the platform
- Being anonymous to platforms and other users

The 'disposability' of an account seemed to have an impact on a user's behaviour, and should be explored alongside anonymity and abuse.

Highly disposable accounts – often called 'burner' accounts – with few or no followers or previous posts were often cited by respondents as a source from which other users could be abused. If the disposable account was banned, its creator could set up a new account with no significant repercussions or penalties associated with their behaviour.

The benefits to removing anonymity could be achieved without removing anonymity

Few perpetrators of abuse told us they would no longer behave abusively if the ability to be anonymous was restricted.

We have seen in this research that removing the ability to be anonymous online would likely increase the friction involved in creating accounts online, make it more difficult to set up or post from disposable accounts and make it easier in some cases for platforms to enforce their community guidelines. Yet there are other ways in which these benefits could be achieved, which do not involve restricting anonymity online.

Removing the ability to be anonymous online would likely have a negative impact on a wide range of people

There are good reasons people might want to be anonymous online, which are not linked to perpetrating or wanting to perpetrate abusive behaviour. For example, they might use anonymity for their personal safety, or to explore their sexual identity.

Some respondents told researchers that removing or reducing the ability to be anonymous on social media platforms would have a significant negative impact on them. An inability to be anonymous on platforms could mean some vulnerable people are excluded from parts of the internet.

However, people who benefited from anonymity were largely benefiting from being anonymous to other users – most would accept being identifiable to the platform itself. A move to reduce anonymity by stipulating that users must be identifiable to the platform would likely have a lesser negative impact than insisting that all users should be identifiable to other users. Many individuals benefiting from the protections of anonymity told us they were concerned with being anonymous to other users, not the platform.

Limiting the disposability of accounts and/or insisting that users should be identifiable to the platform, could be potential solutions

Given that victims receive abuse from accounts that are anonymous to varying degrees, restricting the ability to be anonymous online might prevent some, but not all, online abuse.

However, the evidence from this research found that much of the abuse victims received from anonymous users was from disposable 'burner' accounts. Rather than removing the ability across the board for people to post from anonymous accounts, it seems likely that making it more difficult to set up and post from these burner accounts specifically could significantly reduce the amount of abuse.

For example, platforms could make it more difficult to set up new accounts, or put a time delay on posting from them, or require more information from their creators so that even if other users could not see who they were, the platform had that information.

Anonymity was often an outcome of abusive behaviour, not a precursor

Some of the perpetrators of abuse interviewed for this research, and some of the experts we interviewed about online behaviour described a 'disinhibition effect' of being anonymous online – that led to users sometimes feeling they could be more aggressive or unpleasant online or seeing others behave in that way. It appears that online anonymity can be a facilitator of abuse.

However, there is limited evidence in previous work, or from this research, that proves that anonymity on its own drives or worsens abuse.

Some users who post abuse actively want to remain identifiable. For example, they feel their posts and the reaction to them brings them status, or they believe what they are posting is 'the truth' and does not constitute abuse.

In this research, we interviewed some users who had not initially been anonymous online, and had not wanted to be anonymous, but who had chosen to become anonymous after they had been banned from online platforms under their real names. Sometimes these users described feeling 'forced' to become anonymous. In these cases anonymity was not driving the behaviour, but was an outcome of that behaviour being sanctioned. We describe such users' 'journey to anonymity' in the body of the report.

Platforms' design, combined with users foregrounding their online identities as victims, can create incentives for cycles of abuse.

The reward mechanisms that are designed into social media platforms – likes, follower counts, etc – incentivise certain behaviours. In relation to abuse, both perpetrators and victims of abuse in this research could be seen to be doing things that were arguably playing a part in perpetuating abusive cycles of behaviour.

Some of the perpetrators of abuse were incentivised to post shocking or inflammatory content to get attention, and were rewarded with 'likes', other people sharing their content or more followers.

Some of the victims often shared the abuse they received, to 'call out' the abusers or fight back. This led in many cases to an outpouring of support and attention. The victim could see this as a positive outcome, but it does also provide a larger audience for the abuser, and could incentivise those who are abusive in pursuit of attention to repeat, copy or escalate their behaviour.

Other respondents who were victims of abuse also acknowledged this cycle, saying they had chosen not to share the abuse precisely because they did not want to do anything that might perpetuate it.
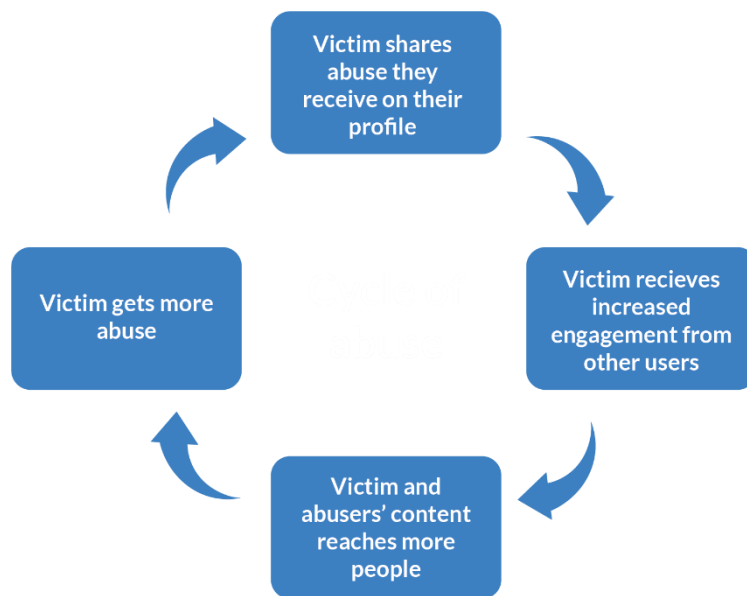

Figure 1: Diagram showing cycle of abuse

Section 1
Introduction

The question of where the balance should lie between allowing anonymity online and protecting people from abuse plays into a wider ongoing debate over the balance between freedom of speech and protection online.

Each perspective comes with significant lobbying power, strong and valid concerns, and a range of conflicting policy demands.

For some, concern over online harm warrants increased identification, verification, and accountability for internet users. An element of this debate relates to online abuse. Some argue that anonymity creates the conditions for abusive behaviour – and there is some evidence to support this view[3].

For others, the trade-offs associated with restrictions to anonymity online are deemed unacceptable. The protection of privacy, freedom of speech and creative expression are often used as arguments for why degrees of anonymity must be protected. Those on this side of the debate also present evidence to suggest reducing anonymity would not reduce online abuse[4].

This research aimed to facilitate DCMS's understanding of arguments for and against anonymity online, and to collate evidence to support their role in informing and improving online safety. This study has provided insight into the behaviour and motivations of people who choose to be anonymous online, those who are abused, and people who abuse others.

The Online Safety Bill[5] will require services in scope to take robust action on anonymous abuse online. At the time of writing, all companies will be required to assess the functionality of anonymous and pseudonymous profiles, and the role they play in allowing illegal content and, for category 1 companies and services which are likely to be accessed by children, legal but harmful content to spread.

All companies will need to remove illegal content from anonymous accounts. Category 1 companies will have to set out clearly which legal but harmful content they do or do not accept on their platform for adult users. All companies will have to protect children from harmful or inappropriate content, including from anonymous accounts.

In addition, major platforms will also have to provide all adult users with the option to verify their identity and give them control over who they interact with. This will help provide robust protections for adults, including vulnerable adults, and allow users to have more control over their online experience.

The aims and focus of this research

Progress on the issue of online abuse is far more complicated than answering the question: would limiting anonymity lead to a reduction in online abuse? Instead, all sides of the debate needed to be explored in this research to support an understanding of how to reduce abuse and its harmful impacts, without disproportionately reducing the benefits of unrestricted, anonymous online participation.

This research sought to understand the relationship between different degrees of anonymity and online behaviour – both the positives (expression, exploration, protection, debate) and the negatives (abuse, harassment, bullying).

Specific objectives of the research were to understand:

---

[3] Babbs, D., 2019. *Clean Up the Internet*. [online]

[4] Bennett, A. and Beverton-Palmer, M., 2021. *Social Media Futures: How to Reconcile Anonymity, Abuse and Identity Online.* [online] Institute for Global Change.

[5] Department for Digital, Culture, Media and Sport., 2022. *Bill 285 2021-22 (as introduced): Online Safety Bill.* House of Commons.

- What is the relationship between different degrees of anonymity and online abuse?
- How does anonymous abuse affect victims?
- What other factors beyond anonymity shape abusive behaviour?
- How do some individuals and groups benefit from anonymity?
- What are the implications of reducing the ability to be anonymous online?

Methodology

This section describes the three parts of the research – a rapid evidence review, expert interviews and in-depth interviews with victims and perpetrators of abuse, and those who benefit or rely on anonymity.

Rapid evidence review

We began by conducting a rapid evidence review of research and literature focusing on abuse, anonymity and the links between the two, understanding the evidence of correlation between abuse and anonymity and any gaps in knowledge or evidence.

Specifically, the review covered:

Definitions of anonymity

Understanding how anonymity is defined both within academic literature and by platforms. Including how anonymity functions on social media.

Existing research into the role anonymity plays in facilitating online abuse, and any evidence of causal links

A review of studies on the relationship between anonymity and online abuse.

Key findings from each source were reviewed, as well as an assessment of the validity of each source. From this, overall reflections were provided on the arguments and evidence put forward for and against a link between anonymity and abuse.

International approaches to dealing with anonymity and abuse

Available policies or draft bills on anonymity and abuse online from a range of countries were reviewed and summarised. Countries covered include:

- Germany
- Australia
- China
- South Korea
- Japan
- Brazil
- France
- Canada

The benefits of anonymity for certain user groups

A range of benefits to being anonymous online were explored. Much of the literature review was based on secondary research into related topics, for example journalistic sources on the impact of any move to reduce anonymity.

Existing verification methods and measures to deal with online abuse by platforms

Publicly available information, provided by platforms, was used to develop a summary of their approaches to dealing with online abuse.

Approach to the review

This was not a systematic literature review and did not cover all existing evidence relating to the link between abuse and anonymity.

The studies reviewed for this report were largely published in scientific journals, meaning they have undergone some form of peer review.

A comprehensive list of search terms were not recorded as this was a rapid review, however, search terms broadly covered:

Online abuse				
Abuse	Online abuse definition	Online abuse prevalence	Recipients of online abuse	Cyberhate
Trolling	Online hate	Online harms	Impact of online abuse	Victims of online abuse
Online anonymity				
Anonymity	Online anonymity definition	Platforms and anonymity	Social media and anonymity	Benefits of anonymity
A search for social media platforms' approaches to anonymity				
The relationship between anonymity and abuse				
Anonymous versus known abuse	Link between anonymity and abuse	What factors lead to online abuse	The Disinhibition Effect	The impact of anonymity online
Anonymity and social media	Cyberbullying and anonymity	Online abuse and anonymity	Online hate and anonymity	
Existing policy recommendations				
Online Safety Bill call for evidence	Anonymity and abuse online	Anonymity and abuse online policy recommendations	*A search of charities and lobbying organisations in this space*	
International approaches to dealing with anonymity and abuse				
Online anonymity legislation	International approaches to anonymity	International approaches to abuse	International online safety	*A search for countries known to have put in place measures in this space*

Searches took place on Google Scholar, EBSCO Discovery Service for academic literature, and Google Search for journalistic articles, research commissioned by charities, and policy recommendations.

In addition to using search terms to find literature, much of the literature explored was found through citations and links from other papers, i.e. exploring literature cited across multiple evidence reviews.

Expert interviews

We then discussed the perspectives of **13** social media platforms and relevant stakeholder organisations that focus on topics such as online safety, abuse and anonymity.

The organisations included:

- **4 x social media platforms**
- **4 x organisations supporting victims of online abuse**
- **5 x think tanks, research or campaigning organisations that have perspectives on online abuse and anonymity**

These interviews took the form of video interviews with representatives from these platforms and organisations, lasting up to one hour. The interviews were semi-structured and covered a range of topics relevant to each organisation:

- The type of abuse they have observed, and who is most affected
- The role anonymity plays in online abuse, as well as other drivers of abuse
- Benefits of anonymity for certain users
- Implications of restricting anonymity
- Other policy changes or measures they believed would reduce the prevalence and/or impact of online abuse.

In-depth interviews

The central part of the research consisted of speaking to victims of abuse and perpetrators of abuse, as well as those who valued anonymity online. Qualitative in-depth interviews were conducted face to face (usually in the home or a neutral space) or via video call, and for two people who preferred to remain completely anonymous to the researcher, via written chat. Interviews lasted up to two hours. Interviews were semi-structured, using a topic guide to shape the conversation.

The decision primarily to use a face-to-face method for interviews was taken to better allow for:

- The sharing of sensitive information and an understanding of the impact of abuse. Receiving abuse can be an extremely personal and distressing experience, which many may not be willing to disclose in a group scenario.
- Discussion about socially undesirable behaviours. Researchers needed to build trust with those who had been abusive online, to be able to discuss behaviours which may be seen as socially undesirable.
- Enabling the collection of objective data on abusive experiences, by asking respondents to show researchers their online profiles and content (abusive and non-abusive) they have been exposed to or have posted on different platforms.
- Going beyond initial reactions to online identification. People had strong reactions to this topic – during the in-depth interviews both initial and more considered reactions were uncovered.

Sampling and recruitment for the interviews

The sample for the in-depth interviews was shaped by the evidence review findings, as well as discussions with DCMS. A total of **23 individuals took part**. There was overlap between groups in the sample (e.g., some perpetrators who were also victims), which is shown in the Venn diagram in *Figure 2*. The sample therefore covered:

Perpetrators of online abuse x 9. Those who had behaved in 'abusive' ways online, including those who did not themselves classify their behaviour as abusive but had been suspended or banned from platforms due to breaching community standards.

Victims of online abuse x 18. People who had received personally targeted online abuse. Given the existing research and discussion around abuse targeted at high profile individuals, such as MPs, we recommended a focus on non-high-profile individuals in this research.

Those who experience the benefits of anonymity online x 10. Those who rely on anonymity to enable them to do what they want to do online. For example, some domestic abuse survivors, people who are LGBT+, undocumented migrants or asylum seekers, and users with 'double profiles' such as teachers or sex workers.

Note: the research set out to include an even spread of victims and those who benefit from anonymity, with a skew towards perpetrators of online abuse. However, the final sample is weighted towards victims of abuse. This was partly due to the difficulty in recruiting those who had perpetrated abuse online, as well as the fact that many perpetrators and those who benefit from anonymity had *also* been victims of abuse.

Across the sample, we aimed to include a mixture of ages, gender, ethnicity, sexuality, socio-economic background. We also ensured a range of platforms were covered within the sample.

The final sample included users of a range of social media platforms including Facebook, Instagram, Twitter, YouTube, TikTok, Snapchat, Gab, Gettr, BitChute, Rumble, and Tattle Life.

Respondents were accessed through in-house recruitment conducted by the research team. This involved:

- Spending time on different social media platforms – identifying and contacting individuals who were victims, perpetrators or assumed benefiters from anonymity. This involved posting in a variety of groups and threads to advertise the project. Researchers also searched for evidence of online abuse across the range of social media platforms mentioned above, by looking through threads, groups, pages, and individual accounts who were speaking about topics likely to spark discussions or debates which may lead to abusive behaviour. For example, discussions about immigration, transgender rights, politics and religion. This was not a systematic review or search for online abuse, but rather one way to find potential respondents.
- Paid-for adverts on social media platforms
- Working with organisations who support victims of online abuse or those who may benefit from anonymity to share information about the research with their clients or members

The research was framed in different ways to reach different audiences. For example, adverts / posts which were aimed at victims asked questions such as 'have you been a victim of online abuse?', while adverts / posts aimed at perpetrators asked question such as 'have you been suspended from social media platforms?'.

Once respondents had expressed an interest in the research, they were given an information sheet explaining that this research aimed to explore the link between anonymity and abusive behaviour online. It stated that we were speaking to a wide range of people with different online experiences to capture a balanced picture of how online abuse and anonymity were experienced in the UK.

Ethical considerations

Steps were taken to ensure the research was ethical and respondent wellbeing was considered at every stage.

- All respondents were given an information sheet before agreeing to take part in an interview, outlining the purpose of the research, what taking part would involve, and how their information would be used.
- The research was entirely optional and potential respondents were reminded that they did not have to take part if they did not wish to
- Respondents were given the option to take part in interviews in a way they felt most comfortable with – whether this was an in-person interview, via a video or telephone call, or via anonymous online instant messaging
- Respondents were reminded at the start of the interview that they did not need to answer any questions they did not feel comfortable answering, and could stop the interview at any time
- All researchers carrying out interviews were experienced in speaking to potentially vulnerable individuals about sensitive subjects, and safeguarding processes were in place, should this have been necessary
- Respondents were given the opportunity to ask questions about the research before giving consent for their information to be used. Respondents were provided with contact details of a researcher should they wish to get in touch after taking part in the research
- All respondents were given pseudonyms for use in this report, and identifiable details such as their locations or job roles were changed or removed to protect their anonymity

Approach to analysis

During the qualitative analysis for this project, respondents' experiences were explored in detail.

The team compared experiences across a number of themes and research questions, including:

- What different types of abuse and anonymity online look like
- How and whether experiences of abuse were linked to anonymity
- How anonymous abuse and non-anonymous abuse affected victims
- How other factors beyond anonymity were shaping abusive behaviour
- How some individuals were benefiting from anonymity.

An analysis grid was created, detailing the findings from each respondent against each of the themes and research questions listed above. This enabled the team to:

- Create frameworks. For example, a categorising the types of abuse observed during this research.
- Map respondents onto different spectrums. For example, their level of anonymity online (i.e. from fully anonymous to fully identifiable), the 'disposability' of their social media accounts (the extent to which their social media accounts were stable vs unstable), and the impact of anonymity on their online experiences (mapping the positive and negative outcomes which they attribute to anonymity online).
- Explore similarities and differences across respondents, and spot patterns. For example, explore whether attitudes towards the ability to be anonymous online differed for those who were victims of abuse compared to those who had perpetrated abuse online.

In conducting the research, we identified different instances of abusive behaviour online. During interviews, victims showed researchers screenshots, photos and live examples of the abuse they had received, while perpetrators showed researchers content they had shared or sent. Researchers also observed some abusive content and behaviour while searching online platforms to find potential respondents to take part in the research. Specific instances of abusive behaviour were analysed to identify the level of anonymity associated with the account it came from (where possible to infer this), and the type of abusive behaviour.

Limitations of the approach

Methodology

As a small qualitative study, this research cannot be used to quantify experiences or definitively prove or disprove the link between anonymity and abuse.

As there are a number of factors which influence behaviours online – including specific platform design features and interactions with and between other users – isolating the impact of anonymity on abusive behaviour is particularly difficult.

Instead, this qualitative study was able to provide in-depth insight into experiences, people and their behaviours – delivering a detailed understanding of the nuances and consideration of the factors at play within this complex topic.

Sampling and recruitment

While efforts were made to achieve a sample with a range of experiences and perspectives, some types of people were harder to reach than others.

For example, while the team did speak to anonymous perpetrators of abuse, some perpetrators exhibit more malicious behaviours than those we were able to speak to. Individuals that are extremely protective of their privacy, and in some cases, are posting abusive content that crosses into an illegal space, are highly unlikely to agree to take part in research.

To identify participants for the research, one of the methods used involved researchers looking for evidence of abusive behaviour online and then contacting the perpetrator of the abuse. One challenge encountered was that abusive content appeared to be removed relatively quickly from platforms, along with perpetrators of abuse having their accounts suspended. While this action is positive and welcomed by victims, it meant it was more difficult to identify perpetrators using this recruitment method.

Note:
This report contains themes which may be offensive or upsetting

Given the subject of this report, some respondents and experts described content and themes which some readers may find offensive or upsetting. This includes:

- Explicit language

- Content relating to self-harm or suicide
- Content of a sexual nature
- Racially offensive language

The respondents

As described above, we interviewed people who benefited from the protections of anonymity, who were victims of abuse and who were perpetrators of abuse.

The research highlighted significant overlap across these groups. For example, we spoke to people receiving abuse who were also perpetrating it, and people who benefited from anonymity themselves but also faced abuse from other anonymous accounts.

We have structured part of the report (sections four, five and six) by respondent groups, to ensure we fully explore each perspective. However, overlap between experiences means we often refer to the same respondent in more than one section.

The diagram below shows the categories each of the people we spoke to fit into. **All names used are pseudonyms** and certain details have been changed to protect the anonymity of those taking part (e.g. ages and job roles).
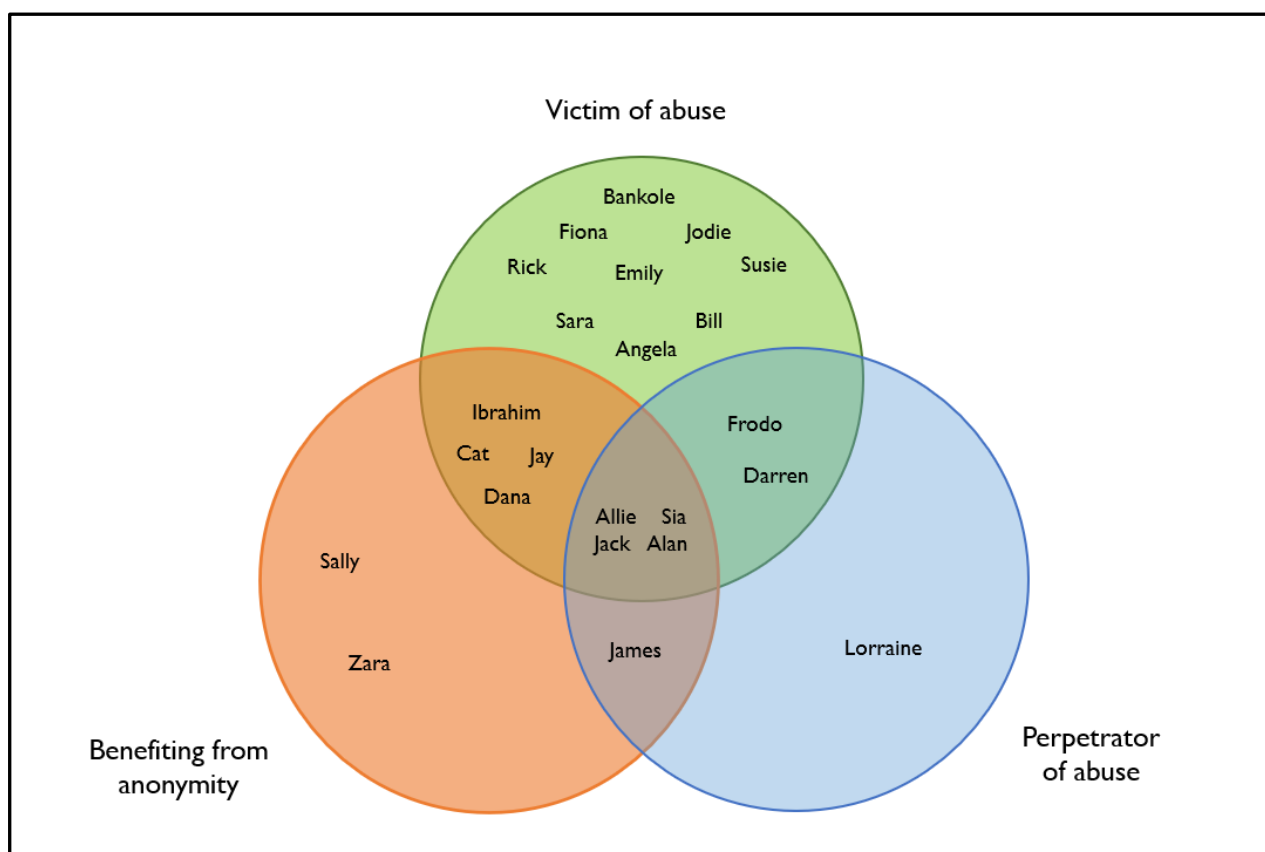

Figure 2: Venn diagram showing where respondents fit in relation to the three categories of victims, perpetrators and those who benefit from anonymity

Lorraine is 56. She is white British and lives on her own. She speaks a lot online about her anti-immigration views. She has been banned multiple times from Twitter and now spends most of her online time on Gab.

Dana is 28. She moved from Scotland to England, where she now works in a pub. She uses largely social media to keep in touch with friends, share her art and read about politics. She was the victim of a pile-on by a group of boys from her hometown when she got involved in an argument about Covid regulations and racism.

Bankole is 28. He is black British, lives alone and works from home as a programmer. He uses Twitter a lot to engage in debates about politics, religion and history but these tend to be very civil. When he was at university, he was the victim of a pile-on from other students for his opinions he posted in a student group about immigration policy.

Bill is 27. They are white British and live in a rural area where they grew up. They are a writer who often speaks about gender and sexuality and have a large following on Instagram. They have received a range of abuse – from abusive comments, to pile-ons, to threats of violence.

Ibrahim is 22. He lives with his parents and has just graduated university. He spends his days gaming, job searching, going to the gym and looking at social media. He gets into arguments online about history, politics and cryptocurrency, a few of which have included abusive language. He was bullied at school and kept his accounts anonymous for a long time because he felt that when getting into heated arguments online, people will always go to your profile and make comments about the way you look.

Angela is 47. Three years ago, she came out as a trans woman and is undergoing surgical transition. Angela started a TikTok account during the pandemic, where she spoke about being trans. She has received a range of abuse: harassment, pile-ons, impersonation, doxxing and physical threats.

Jodie is 24. She is mixed race with white British and black Caribbean parents, and lives on her own. Jodie runs an Instagram blog focused on self-love, anti-racism and body positivity, where she often posts about social justice issues and her experience, as she describes it, as a black and queer woman. She has experienced a variety of abuse, mostly reactive, angry comments about her weight and race.

Fiona is 25. They are white British and live with their family. Fiona has a large social media platform, particularly on TikTok. Fiona identifies as non-binary and much of their online content is related to issues of social justice for the LGBTQ+ community, mental health and body positivity. Fiona describes themselves as "fat" and receives a lot of sustained abuse about their appearance and gender, ranging from 'flippant' comments to more serious intimidation encouraging suicide and self-harm.

Emily is 30. She lives in with her partner. Emily runs a large Instagram blog which talks about women's issues and feminism. Emily has received abuse via this blog since its inception - typically misogynistic or agenda-driven comments and direct messages.

Allie is 33. She co-parents her two children who stay with her regularly. Allie is transgender, and recently transitioned socially. While Allie has experienced online abuse in relation to her gender identity, she has also benefited greatly from anonymity online. For Allie, anonymity provided a means to find support in her transition and explore her identity with freedom and protection from the gaze of offline family and friends, before she was ready to transition socially.

Sara is 56. She was in an abusive relationship and relies on anonymity online to avoid detection and harassment from her ex-partner.

Sally is 30. She lives in with her partner. Sally left a fundamentalist religious group in recent years and has used anonymity as a means to exist freely online; to share her political views and opinions, to get support from other people who have left a fundamentalist community, to get health advice and to use social media without detection from her work or her previous community.

James is 70. He lives with his partner and granddaughter. He identifies as 'far right' and has anti-immigration and anti-vaccination views. He engages in debates online about religion. His Twitter account was blocked in January 2021, though he was not sure why, and since then he has been using Gab, Gettr, BitChute and Rumble.

Fred is in his mid-20s and works as a security guard. He spends a lot of his spare time online, Twitter being his main platform. Fred self-identifies as a "troll" and spends a lot of time deliberately "baiting" people who he thinks are contributing to an overly "censorious" environment. He has been suspended and banned from Twitter multiple times but finds it incredibly easy to "respawn" – setting up new accounts with a different pseudonym each time.

Alan is 60 and lives with his wife. He has a keen interest in politics and started using Twitter while he was living in England in order to keep up with the debate about the Scottish referendum. Alan noticed a change over time in how debate had changed on Twitter, becoming "over-sensitive, identity-based and censorious". After being temporarily suspended from Twitter, Alan decided to leave the platform and migrate to Gettr, where he felt debate was less stifled.

Jack is 19. He is mixed race and comes from a devout Muslim family on his dad's side. Jack writes screenplays and scripts, does some work in radio and aspires to work in media. He describes himself as left-leaning and passionate about politics. Jack is anonymous on Reddit and Twitter because he wants to be involved in debates on topics such as race, gender and sexuality without his own identity coming into it.

Jay is 21, mixed race and in his third year at university. He spends a lot of time on his own, mostly gaming, watching YouTube and on social media. Jay used to get into debates on the topic of race and would experience either 'implicit' or 'explicit' racism during these debates.

Rick is 44. He has two daughters and is passionate about local politics – which takes up the majority of his time. He also has "severe dyslexia", and said he is almost illiterate as a result. Rick is passionate about having open and public debates about politics. As a result, he has received a lot of abuse on social media, which sometimes starts being political, but often descends into personal attacks – for example, commenting on his ability to be a father and saying he deserved to be left by his partner.

Susie is 22 and lives with her partner. Susie made a decision at work which caused some controversy among those the decision would affect, and she started to receive abuse as a result. The majority of the abuse was posted anonymously on anonymous confessions pages, where an anonymous admin would collate and post all the posts on the page. She felt uneasy at not knowing where the abuse was coming from.

Darren is in his late thirties. He practices almost absolute anonymity online, using technology such as VPNs and alternative browsers to increase his ability to use the internet and not be traced. Darren is active on Gab and Reddit, using pseudonyms and revealing no details about his life. He sometimes posts inflammatory statements and abuse online.

Zara is 24. She escaped a forced marriage in Pakistan when she was 21 and has since used a pseudonym on social media and dating apps to avoid being identified by her old family and network of family friends. As a single mum who is relatively socially isolated, Zara relies a lot on things like Facebook Marketplace and various social media groups for support – and feels quite strongly that she would be effectively excluded from much of online life if she were forced to use her real name.

Cat is 24. She is an adult content creator and works hard to maintain a division between her work identity and her personal identity. In the past, some people have "worked out" who Cat "really is" and threatened to contact her family. She said she would not be able to do her job if she had to use her real name.

Sia is 25. They identify as non-binary. Sia has strong left-wing political views and opinions, particularly in relation to race and gender. Sia is often abusive online, arguing with and ridiculing those with different political opinions.

Section 2
Evidence review summary

This summary outlines learnings from the rapid evidence review, which set out to explore:

- Definitions of abuse and anonymity
- The benefits of anonymity for certain user groups
- Existing verification methods and measures to deal with online abuse by platforms
- Existing research into the role anonymity plays in facilitating online abuse, and causal links
- International approaches to dealing with anonymity and abuse

The summary includes an overview of all the sections within the review. The full review can be found in an annex to this report and includes thorough analysis and detailed tables of all the evidence explored.

Online abuse

Defining online abuse

Online abuse is a broad category containing many different types of behaviour, which are likely to have different outcomes depending on the severity, type and recipient. Definitions of online abuse are highly subjective and dependent on interpretations of intent and impact. Imprecise definitions of abuse can lead to discussions of online behaviour that conflate incivility or impoliteness with abusiveness[6] [7].

We have included explanations and definitions of abuse and anonymity in section 3 of this report.

How widespread is online abuse?

The prevalence of abuse online has predominantly been measured using **surveys** and **measurement studies** (in which academics analyse samples of social media data e.g. a dataset of Tweets, to detect the prevalence of potentially harmful content[8]). **Platform transparency reports** have also been used to indicate prevalence of online abuse.

The limitations of these methods for measuring the prevalence of abuse have been explored in detail within DCMS's Online Harms Feasibility Study[9]. The study found that while a variety of measures and sources exist, they do not provide an accurate or consistent enough picture of online harm to measure reliably the prevalence or impact over time — there is no 'ready-made' solution — and there are numerous challenges to measuring online harm well.

Which users are more likely to experience online abuse?

There is an extensive amount of literature that explores whether certain groups are more likely to receive online abuse than others. This literature has found that:

- **Ethnicity:** From a review of UK studies on online abuse, those from Black and 'Other' ethnicities were more likely to report being targeted with online abuse compared with White or Asian people[10].

---

[6] Papacharissi, Z. 2004. 'Democracy online: civility, politeness and the democratic potential of online discussion groups', *New Media & Society* 6(2):259-283

[7] Muddiman, A. 2017. Personal and Public Levels of Political Incivility. *International Journal of Communication* 11

[8] DCMS, 2021. *Online Harms Feasibility Study*. Unpublished (to be published in 2022)

[9] Ibid.

[10] Vidgen, B., Margetts, H. and Harris, A., 2019. *How much online abuse is there? A systematic review of evidence for the UK Policy Briefing — Full Report.* Public Policy Programme Hate Speech: Measures and Counter Measures. [online]. The Alan Turing Institute.

- **Sexuality:** Lesbian, gay or bisexual adults in the US were more likely to report receiving online harassment compared with straight adults (51% of Lesbian, Gay and Bisexual (LGB) adults targeted with more severe forms of online abuse compared to 23% of straight adults)[11].
- **Disabilities:** From a review of UK studies of online abuse, those with disabilities reported seeing more online abuse than those without disabilities[12].
- **Gender:** A review of studies in Scotland and the UK exploring the prevalence of online abuse for trans individuals found that 34 to 52% of trans people had experienced online abuse[13].
- **High profile individuals:** Qualitative research with UK MPs[14], as well as measurement studies[15] have highlighted the online abuse received by MPs. Other high-profile individuals such as footballers and influencers have also been highlighted as key targets for online abuse[16] [17].

However, few of the existing studies of UK adults explored in the examples above cover a nationally representative population. Instead, they focus on specific groups (e.g. experiences of MPs). This means it is not possible to breakdown data relating to experiences of online abuse across different demographic groups, making it hard to compare prevalence rates among UK adults.

Impact of online abuse

Some literature has highlighted the impact online abuse can have on[18] [19] [20]:

- **Emotional and mental health.** For example, people who have experienced online abuse report suffering from anxiety, depression, and panic attacks. Other reports include lower self-esteem, and feelings of anger and shame.
- **Fears over physical safety.** People report fearing for their physical safety offline.
- **Changes to online activity**. For example, people who restricted use of their social media account, or deleted their accounts altogether.
- **Impacts on friends and family**. Friends and families of victims also report being affected – often becoming concerned about the victim's safety, as well as being targeted or threatened themselves.

However, it is worth noting that many studies exploring the impact of online abuse do not specify the *type* of online abuse someone has received and explore how this relates to its impact. The impact of abuse can also vary depending on the recipient.

Is anonymous abuse experienced differently by victims from 'known' abuse?

There appears to be limited research on the impact of known versus anonymous abuse on victims. In a qualitative study with 25 young people aged 15 to 24 on the risk factors associated with the impact and severity of online bullying, participants reflected on what impact the anonymity of the perpetrator had on feelings of loneliness, fear and worry[21]. Some respondents felt increased impact if their bully was anonymous as

---

[11] Pew Research Center: Internet, Science & Tech. 2021. *The State of Online Harassment.* [online]

[12] Vidgen, B., Margetts, H. and Harris, A., 2019. *How much online abuse is there? A systematic review of evidence for the UK Policy Briefing – Full Report.* Public Policy Programme Hate Speech: Measures and Counter Measures. [online] The Alan Turing Institute.

[13] Ibid.

[14] Committee on Standards in Public Life, 2021. *Intimidation in Public Life: A Review by the Committee on Standards in Public Life.* [online]

[15] Farrell, T., Bakir, M. and Bontcheva, K., 2021. *MP Twitter Engagement and Abuse Post-first COVID-19 Lockdown in the UK: White Paper.* [online]

[16] UK Parliament Committees. 2020. *Anton Ferdinand, Lianne Sanderson and Marvin Sordell give evidence on online racist abuse.* [online]

[17] BBC News. 2021. *Social media influencers face relentless abuse, MPs are told.* [online]

[18] Hubbard, L., 2020. *Online Hate Crime Report: Challenging online homophobia, biphobia and transphobia.* Galop.

[19] End Violence Against Women and Glitch, 2020. *The Ripple Effect: Covid-19 and the Epidemic of Online Abuse.*

[20] Amnesty International. 2017. *Amnesty reveals alarming impact of online abuse against women.* [online]

[21] Dredge, R., Gleeson, J. and de la Piedad Garcia, X., 2014. *Risk Factors Associated with Impact Severity of Cyberbullying Victimization: A Qualitative Study of Adolescent Online Social Networking. Cyberpsychology, Behaviour, and Social Networking, 17(5), pp.287-291.*

there was a fear in the unknown: *"I felt victimized and scared because I did not know who'd done it"*. However, others felt more impact when their bully was known to them: *"That had a huge effect on me because that person used to mean so much to me."* This suggests that while type of impact may be different, it is difficult to determine whether anonymous abuse has more impact on the receiver: this will vary greatly depending on the individual, context and content of the abuse.

Methods platforms use to combat online abuse

Social media platforms currently use a range of tools intended to combat online abuse[22] [23] [24] [25]. These fall into three broad categories:

- setting rules and guidelines about online abuse
- identifying online abuse
- acting on online abuse

The full evidence review includes a table illustrating the tools used by platforms under these three categories and a description of how they operate.

Online anonymity

Online anonymity cannot be understood in binary terms. There are different degrees of anonymity, which can have different implications for users and should not be conflated. For example, a platform could allow a user to be entirely anonymous – unidentifiable to both the platform and to other users – or they might require the user to be identifiable to the platform but allowed to use a pseudonym – concealing their identity from other users.

Anonymity is used for a range of reasons. By freeing people from the social norms and reputational risk that guide their behaviour *offline*, anonymity might facilitate people posting abusive content, but it might also be used by people who want to be themselves online in a way they feel they cannot be offline. The versatility of anonymity is precisely what makes policy decisions so complex.

Platforms and anonymity

Most popular social media sites enable a level of pseudonymity, except for Facebook and LinkedIn which encourage 'real name' policies. However, these policies are difficult to enforce, and have received criticism for cultural bias[26]. Moreover, they are often not able to determine that the *real-sounding* name you use is actually *your* name.

A table outlining platforms' approaches to anonymity is included in the full evidence review.

Who benefits from anonymity?

Discussions regarding the benefits of anonymity online for particular people or groups of people are largely anecdotal or based on secondary research into related topics. It focuses on those who, due to their offline identities, occupations, social networks, or communities, may be at risk both online and offline if their identities were not concealed. Much of this conversation in the UK was inspired by the emergence of the idea of introducing real-name policies and ID verification across social media platforms following the racist abuse of prominent footballers during the Euro 2020 tournament[27]. It was argued that reduced anonymity online by the

---

[22] Facebook Meta, 2022. *Transparency Center.* [online]

[23] Twitter. 2022. *Help Center.* [online]

[24] TikTok. 2022. *Safety & privacy controls | TikTok.* [online]

[25] Reddit. 2022. *Content Policy - Reddit.* [online]

[26] Herrman, J., 2021. *You Anon.* [online]. The New York Times.

[27] Smith, A., 2021. *Calls to end social media anonymity give platforms more power without actually fixing the problem, experts say.* [online] Independent.

introduction of real-name policies would exclude those unable to provide ID verification and unfairly expose others[28] [29].

A range of academic papers explore the wider costs and benefits of anonymity online and point theoretically or evidentially, based on small samples, to some groups of beneficiaries.

The relationship between anonymity and abuse

Many reports on online harms are inconclusive about the role of anonymity in abuse, often because they rely on imprecise definitions of *abuse* and *anonymity*.

For example, a 2020 YouGov poll found that 83% of British people think the ability to post anonymously makes people "ruder" online. This has been cited by *Clean Up the Internet* in its arguments that link anonymity with abuse but, as highlighted earlier, it is problematic to conflate abusive behaviour with impoliteness.[30]

Similarly, lots of surveys do not distinguish between *strangers*, *anonymous accounts* and *pseudonymous accounts* when they ask respondents about their experience of online abuse, including those from *The Pew Research Centre*[31], *Amnesty International*[32], *Compassion in Politics*[33] and *Glitch*[34]. Due to the lack of a consistent definition, these surveys cannot be used to definitively demonstrate a relationship between anonymity and abuse.

The links between anonymity and abuse

The studies explored within the evidence review provided mixed conclusions on the link between anonymity and abuse. Many of the studies are limited in their focus, with differing definitions of abusive behaviour and some exploring only a single platform.

Several studies referred to John Suler's theory of the disinhibition effect, which suggests that online communication causes users to lose inhibition and act in ways that they would not offline[35] [36] [37]. This can work in two ways: benign disinhibition (unusual openness, vulnerability, generosity) and toxic disinhibition (rude language, hatred, harsh criticisms)[38]. The principle is that when a person's online persona is not traceable back to their offline self, people feel freer to act differently.

Some of the more recent studies have corroborated this theory. For example, a study using a data-scraping methodology on Twitter that analysed over 100,000 tweets for "extreme" words, found a much higher proportion of "extreme language" from accounts that were anonymous[39].

[28]Bold, B., 2021. *Should social media users be required to prove their ID to stamp out online abuse?.* [online] Campaign.

[29]Dyer, H., 2021. *Online abuse: banning anonymous social media accounts is not the answer.* [online] The Conversation.

[30] Babbs, D., 2019. Clean Up the Internet. [online]

[31] Pew Research Center: Internet, Science & Tech. 2021. *The State of Online Harassment.* [online]

[32] Amnesty International. 2017. *Amnesty reveals alarming impact of online abuse against women.* [online]

[33] Compassion in Politics. 2021. *Three-quarters of those experiencing online abuse say it comes from anonymous accounts.* [online]

[34] End Violence Against Women and Glitch, 2020. *The Ripple Effect: Covid-19 and the Epidemic of Online Abuse.*

[35] Carter, P., Sutch, H. 2019. Anonymity, Membership-Length and Postage Frequency as Predictors of Extremist Language and Behaviour among Twitter Users. *International Journal of Cyber Criminology,* 13(2): 439-459.

[36] Kurek, A., Jose, P. E., Stuart, J. 2019. *'I did it for the LULZ': How the dark personality predicts online disinhibition and aggressive online behaviour in adolescence* [online]

[37] Lapidot-Lefler, N. and Barak, A., 2012. Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior*, 28(2), pp.434-443.

[38] Suler, J. 2004. 'The Online Disinhibition Effect', *CyberPsychology & Behaviour* 7(3) 321-326.

[39] Carter, P., Sutch, H. 2019. Anonymity, Membership-Length and Postage Frequency as Predictors of Extremist Language and Behaviour among Twitter Users. *International Journal of Cyber Criminology,* 13(2): 439-459.

However, there are also studies that have concluded that anonymity is not a driving factor behind abusive or uncivil posts[40] [41]. For example, a large-scale data-scraping study by Rost, Stahel and Frey, which analysed all comments on online petitions published on a socio-political German social media platform, found more online aggression was posted by non-anonymous commenters.

What other factors contribute to online abuse or incivility?

The study by Frey et al. of posts on a German political platform www.openpetition.de, indicates that other users on the platform and the culture of a platform have a significant impact on prevalence and types of abusive content and whether or not anonymity plays a role.[42] The study explored the impact of social norm theory, where 'online firestorms' (collective online aggression directed towards actors of public interest) can enforce social norms by expressing public disapproval.

What could platforms be doing to tackle anonymity and abuse?

An experiment in the r/science community, a sub-community[43]on Reddit, which facilitates discussion of scientific research and had 13.5m subscribers, asked moderators to post the community rules in a 'sticky' comment at the top of some threads[44]. This led to a decrease in users posting in a way which violated community rules. The experiment indicated that interventions by moderators can influence social norms and shape online behaviour and adds to the argument that the culture of a platform has a large impact on the amount of abuse present.

Alfred Moore argues that it is important to separate two aspects of anonymity that are associated with abusive behaviour[45]. The first is a lack of durability: the idea that a user can create temporary anonymous accounts which free them from the reputational pressures of a lasting persona i.e. creating 'burner' accounts to abuse other users. The second is a lack of connectedness: the idea that a user's actions in an online space are not traceable to them offline – they can act differently in offline and online spaces without accountability.

Overall, the above studies suggest that platforms could be:

- Aiming to improve or change the platform culture, to make abuse less appealing to users – ensuring it reduces status, views etc. rather than increasing them.
- Stopping users from creating multiple accounts which means there are no consequences or 'durability'. Currently, if a user is blocked on one account, they are often easily able to create another and carry on with their behaviour.
- More consistently enforcing community guidelines.

Existing policy recommendations

Various policy recommendations to tackle the issue of anonymous abuse have been proposed by different organisations, particularly in reference to the Online Safety Bill. *A detailed table of these recommendations can be found it the full evidence review.* Most recommendations acknowledge the benefits of anonymity online, and therefore do not advocate for an outright ban on anonymity or pseudonymity. Specific proposals are summarised in more detail within the full evidence review within the annex, but broadly they include:

- Explicitly including anonymity as a risk factor for harm (i.e., online abuse) in Ofcom's (the regulator that will be responsible for enforcing the bill) proposed risk assessments for platforms to carry out

---

[40] Lee, G., Sanchez, M. 2018. 'Cyber Bullying Behaviors, Anonymity, and General Strain Theory: A Study of Undergraduate Students at a South Eastern University in the United States' *International Journal of Cyber Criminology, Vol. 12(1): 84–96.*

[41] Rost, K., Stahel, L. and Frey, B., 2016. Digital Social Norm Enforcement: Online Firestorms in Social Media. *PLOS ONE*, 11(6), p.e0155923.

[42] Ibid.

[43] Reddit is a large community made up of thousands of smaller 'communities'. These smaller, sub-communities within Reddit are also known as "subreddits" and are created and moderated by users.

[44] Matias, J., 2019. Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. *Proceedings of the National Academy of Sciences*, 116(20), pp.9785-9789.

[45] Moore, A., 2017. Anonymity, Pseudonymity, and Deliberation: Why Not Everything Should Be Connected. *Journal of Political Philosophy*, 26(2), pp.169-192.

- Increasing friction in the process of creating social media accounts
- Ensuring law enforcement can access sufficient information from social media platforms to tie online abuse to perpetrators when a crime has been committed
- Providing social media users with the option to verify their identity
- Providing social media users with the option to block interactions with unverified users, and enabling users to see whether or not a user is verified

International approaches to dealing with anonymity and abuse

Few countries have so far successfully implemented measures that seek to reduce anonymity to tackle abusive behaviours online. Countries which have existing measures in place have included nations with reduced freedoms for their citizens and a history of censoring content online, such as China[46].

In 2007 South Korea put in place legislation which required the largest Korean-based platforms to collect ID information for people posting on their websites[47]. However, this was overturned in 2012 as it was found to be a disproportionate response to the issue, resulting in data breaches and with limited evidence that it had any impact on online abuse.

Other countries are currently exploring ways to make platforms more responsible for identifying their users, including Australia which is planning to make social media companies collect and store identifiable information[48]. France is also exploring the idea of linking newly released digital ID to social media platforms – which would require users to upload identification upon registration to the platform[49].

All of these example measures still allow users to create profiles which do not require their real name — therefore remaining anonymous to other users. However, the platforms are (or will) be required to collect and store identifiable information on users which will increase accountability for their actions online.

Conclusions

The links between abuse and anonymity are widely contested across the studies explored within this rapid evidence review. However, many studies are limited in their methodology or scope — with variations in definitions of abuse, limited sampling criteria or a lack of consideration for contextual information and severity of abuse.

Despite this, there are several studies which do suggest anonymity on platforms increases the likelihood of abusive, or uncivil, behaviours, often citing the disinhibition effect. However, removing anonymity is rarely suggested as a suitable solution to reducing abuse, as there are a range of other factors which must be taken into account when assessing the issue:

- Abuse still does occur from non-anonymous users, and there are studies which have found that, on certain platforms, anonymity does not lead to increases in abusive content, and in fact occasionally identifiable accounts are preferred for directing abuse at other individuals.[50]
- Anonymity is highly valued by many. Often individuals experiencing vulnerabilities benefit from being able to remain anonymous online and feel that removing their ability to be anonymous may increase the abuse they personally experience, both online and offline.
- There is a lack of research on the impact of anonymous versus non-anonymous abuse on victims. Within this rapid review, only one study was found which explored impact, where respondents reported mixed

[46] Liao, S., 2017. *China is forcing internet companies to end online anonymity.* [online] The Verge.

[47] Taylor, J., 2021. *Twitter says any move by Australia to ban anonymous accounts would not reduce abuse.* [online] The Guardian.

[48] Healey, B., 2021. *Australia's new Online Safety Act is starting to take shape, with harsh penalties levelled at trolls to include fines of up to $111,000.* [online] The Chainsaw.

[49] Mascellino, A., 2021. *Social media accounts could soon require digital ID in France, UK | Biometric Update.* [online] Biometric Update.

[50] Rost, K., Stahel, L. and Frey, B., 2016. Digital Social Norm Enforcement: Online Firestorms in Social Media. *PLOS ONE*, 11(6), p.e0155923.

views on whether anonymous or non-anonymous abuse was more impactful[51]. This is an important line of inquiry — as resulting harm is an essential factor in deciding what action needs to be taken. For example, if non-anonymous abuse is found to be less prevalent but significantly more harmful, this could have a large impact on the outcome of the debate.

Some studies have suggested other ways to lessen the link between anonymity and abuse, such as ensuring anonymous profiles are at a minimum identifiable to the platform and changing the culture of an online space. According to many studies exploring social norm theory and anonymity, platform culture and the behaviours of other users have a large impact on behaviour, with some platforms or experiments attempting to harness this through changes in platform design and interventions by active and identifiable moderators[52].

Internationally, some countries have taken steps to reduce anonymity online. Often measures are preferred which ensure users are identifiable to platforms rather than other users – meaning they can be personally held accountable for their actions online and referred to law enforcement if required.

This approach reduces the impact on individuals who might suffer from not being able to remain anonymous online. However, there is little data yet on how effective ensuring users are identifiable to the platform is in reducing abuse. In fact, in South Korea, the Constitutional Court found little evidence of such an impact and removed the policy[47].

Overall, there are approaches to limiting anonymity which are likely safe and acceptable for many users. However, until more is known about the definite links between abuse and anonymity, it is not possible to come to a conclusion about whether this will have any positive effect. Similarly, even if much online abuse is currently anonymous, removing anonymity may just lead to an equivalent increase in non-anonymous online abuse.

More research is needed which specifically explores the links between abuse and anonymity – the motivations of anonymous and non-anonymous users posting abuse, and the impact anonymous abuse has on users, in order to inform future action and policy. The qualitative element of this research aims to add insight to this current gap in research.

---

[51] Dredge, R., Gleeson, J. and de la Piedad Garcia, X., 2014. Risk Factors Associated with Impact Severity of Cyberbullying Victimization: A Qualitative Study of Adolescent Online Social Networking. *Cyberpsychology, Behavior, and Social Networking*, 17(5), pp.287-291.

[52] Matias, J., 2019. Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. *Proceedings of the National Academy of Sciences*, 116(20), pp.9785-9789.

Section 3
What do we mean by online abuse and what do we mean by anonymity?

What is online abuse?

Previous research has tended to categorise online abuse as one of two types:

Abuse directed against a group, usually called 'hate speech'. *"[Hate speech] broadly includes negative textual, visual or audio-based rhetoric that attacks, abuses, insults, harasses, intimidates, and incites discrimination or violence against an individual or group due to their race, ethnicity, gender, religion, sexual orientation or disability"*[53]

Abuse directed against an individual, usually called 'harassment' or 'cyberbullying'. A definition for online (and not necessarily illegal) harassment is: *"Aggressive, intentional acts carried out by a group or individual, using electronic forms of contact against an individual"*[54]

In conducting the research, we identified different instances of abusive behaviour online. This abuse was collated from the content that victims had received, that perpetrators had shared or sent, and that researchers had observed while searching online platforms to find abuse. What became obvious was that this behaviour was diverse, and varied greatly in terms of:

● **whom it is directed at** e.g., one person, a group of people or an audience of people who agree with their viewpoint
● **the format/mechanism** e.g., direct messages, posts, images
● **what the intent appears to be** e.g., to hurt or upset someone, to threaten someone, to perpetuate a theory or ideology, to get approval from others
● **the impact it seems to have** e.g., the amount of harm it causes.

Types of abuse observed during this research

Across these variations, it became apparent that there are some distinct 'types' of abuse – different types of abusive behaviour and content. The main types of abusive behaviour and content identified in this research could be mapped in terms of frequency (one-off versus sustained) and target (a person versus a group) as shown in *Figure 3*.

These types of abuse are associated with a range of different behaviours as well as differing levels of anonymity (discussed in more detail later in this chapter). The impact on victims can vary depending on which type of abuse is being directed at them.

---

[53] Davidson, J., Livingstone, S., Jenkins, S., Gekoski, A., Choak, C., Ike, T. and Phillips, K., 2019. *Adult Online Hate, Harassment and Abuse; A Rapid Evidence Assessment.* [online]

[54] Vidgen, B., Margetts, H. and Harris, A., 2019. *How much online abuse is there? A systematic review of evidence for the UK Policy Briefing – Full Report.* Public Policy Programme Hate Speech: Measures and Counter Measures. The Alan Turing Institute. [online]

Person
targeted

Reactive, angry, often 'one-off' behaviour targeted at an individual

Sustained or obsessive behaviour targeted at an individual, closer to **harassment**

Discrediting or 'calling out' individuals

One-off ⟵   ⟶ Sustained

Content appearing to **promote a particular narrative, ideology or agenda**
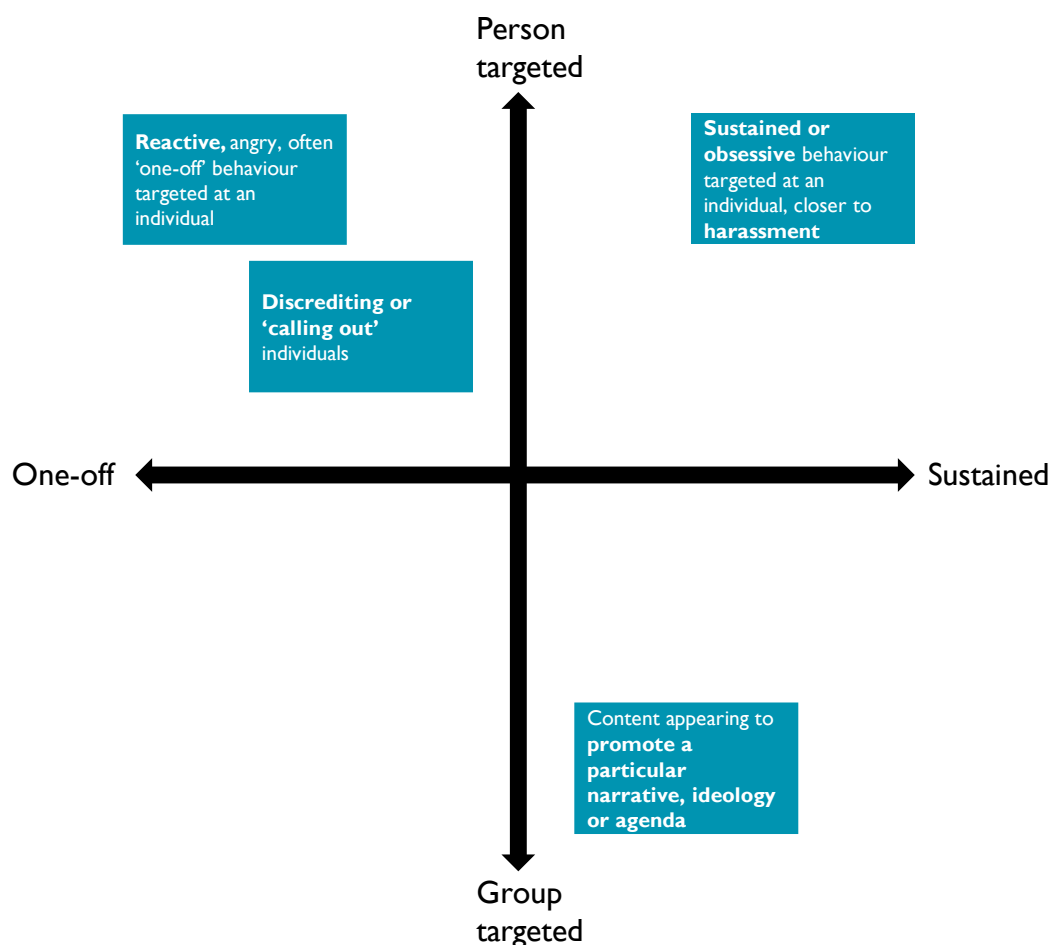
Group
targeted

Figure 3: Types of online abuse observed during this research

Note: this is not a comprehensive model of the different types of abuse which exist online, but summarises the main types of abusive behaviour and content identified during this research.

Some abuse seemed to be sustained or obsessive behaviour targeted at an individual – closer to harassment

This was usually characterised by the same person commenting over and over again on an individual's posts.

Several victims of this kind of abuse identified a number of 'repeat offenders' they were familiar with who would frequently make similar comments on their posts.

One of the clearest examples of this type of abuse came from Fiona's experiences. Fiona is a self-described "fat activist". They regularly post body positivity content on Instagram, and often receive abuse in response. Some of this content was part of an ongoing tirade of abuse against them by individuals or groups of people. One message sent to Fiona said, *"We won't stop until your dead [sic]",* clearly outlining the intention to continue to send them abusive messages.

This kind of abuse sometimes resembles stalker-like tendencies, such as being followed or constantly harassed. For example, two victims of abuse spoke about their address or location being used in threatening messages.

"Someone sent me a pin of where I worked on a map and said they were going to find me and kill me" **Bill, 27**

Similarly, Angela is a trans woman who gained a substantial following on TikTok when she started live streaming videos during the pandemic about being trans and working in the travel industry. Angela subsequently found out there was a page on Tattle – a gossip platform focused on influencers with public social media accounts – that was dedicated to gossiping and talking about her and her partner. Someone on this page had her address and shared it on the Tattle page – which resulted in someone anonymously sending dog faeces to her house.

Other types of abuse appeared to be more reactive, angry, and often 'one-off' abusive behaviour

When victims of online abuse showed us the kind of abuse that gets directed at them, the majority of it could be described as reactive and angry. It was usually a post, comment or direct message posted or sent in reaction to a piece of content posted or shared by the victim.

This kind of abuse is usually characterised by *ad hominem* attacks (e.g. an attack against someone's personality or characteristics, rather than their position), often against someone's appearance.

We interviewed Jodie, who uses her Instagram platform to post body-positivity and anti-racism content. Jodie said that ever since she joined Instagram she had received abuse. This abuse has usually been about the same thing – her weight. Jodie tends to post photos with captions about body positivity, and she gets angry, reactive comments about her size as a response to these photos.

"It's mostly fatphobia and negative comments about my physical appearance. A lot of [people] just seem baffled by the concept that you can be fat and happy. So I get a lot of people saying things like, 'Oh, just die'. Those sorts of things."
Jodie, 24

Cat – an adult content creator – often posts content on TikTok and Instagram in an effort to direct users to her subscription page. When she posts this kind of content, she regularly receives reactive abuse. This includes offensive comments relating to her appearance, intelligence and sexual behaviour.
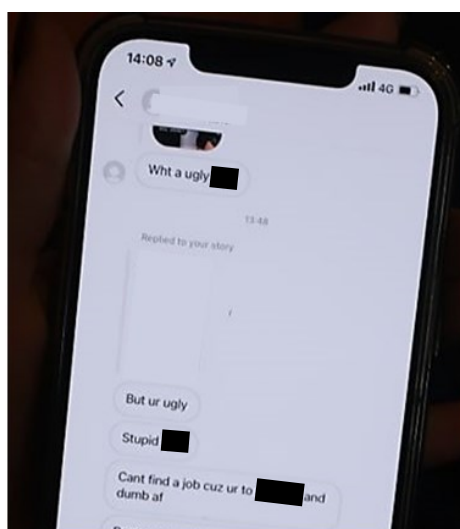

Figure 4: Direct messages sent to Cat, demonstrating 'reactive' abuse.

Some abuse appeared to promote a particular narrative, ideology or agenda

A significant amount of abusive content and behaviour was less often targeted at a specific person but rather seemed to be driven by an effort to promote a certain narrative or agenda.

This content often seems to be perpetuating certain ethnic or religious stereotypes and frequently, though not always, comes in the form of memes.

Sometimes this type of abuse uses or refers to statistics or facts, but it can be inferred to be trying to push a negative narrative. For example, one respondent, Lorraine, is anti-immigration – and is often abusive online while trying to make points about UK immigration.

Some abuse appears to be intended to discredit or 'call out' individuals

This type of abuse seemed intended to discredit or 'call out' individuals whom the perpetrator had determined to be on the wrong side of a cultural argument.

Bill spoke about accounts they saw that were dedicated to "calling people out" for historic tweets or associations that were now seen as "problematic". In their own case, they mentioned that someone had directly messaged them because they followed someone who had written a tweet in the past that was perceived to be antisemitic.

"I remember someone coming into my DMs and saying: 'Why do you follow this person? You support the holocaust. You're like Hitler.'… A bad action on my part could lose me my career. It's terrifying." **Bill, 27**

Bill spoke about there being a culture of this type of abuse, describing the sources as "policing accounts". They suggested that within this community of users, there is a perception that it is good to be extremely critical, bordering on abusive.

Fred, someone who enjoys being provocative and "baiting" people online, described similar accounts, calling them "jannies" – or janitor accounts. These are accounts that go around "getting validation" for calling out unacceptable content and behaviour – often reporting things to the platform.

What do we mean by anonymity?

The word 'anonymity' has several different meanings and applications across our sample. As we set out below, there are differing degrees of anonymity. But the level of disposability of a person's anonymous accounts also has an impact on behaviour.

Degrees of anonymity

There are degrees of anonymity, which can have different implications for users. When reflecting on anonymity without being prompted to do so, most people interpreted the idea of 'being anonymous' to mean **anonymity to other users** of the platform. This requires a pseudonym and usually (although not always) an anonymous picture.

However, this type of anonymity does not mean the user is necessarily anonymous to the platform. Often, when signing up to a platform, a user will need to input some identifiable details such as a real name, email address or telephone number, which will be held by the platform, but not visible to other users. This usually means that in the event of a user committing a crime on the platform, their real identity could be obtained and handed over to the relevant authorities.

A greater degree of anonymity is to be anonymous both to other users and **anonymous to the platform**. The people interviewed who had this degree of anonymity were either on platforms that did not require any identifiable details, such as Gab or Gettr, or they had used false information when setting up an account on a platform that did require those details.

When prompted, some of the wider sample did reflect on the differences between these degrees of anonymity.

Throughout this report, when we refer to examples of abuse, we provide relevant detail of whether the accounts or users involved are anonymous, and to what degree.

Anonymity across different platforms

On most platforms anonymity to other users is accepted and easy to achieve. For example, Instagram, Twitter, TikTok and Snapchat do not require real names to be used when setting up a username that is visible to other users.

Many of these platforms do put in place steps to try to ensure users are identifiable to the platform itself. However, some people interviewed for this research described easily inputting false details. For example, some respondents had signed up to platforms using false names or using emails that had been set up for the sole purpose of signing up to a platform but were not otherwise linked to their identity.

Some platforms have a real-name policy, but do not ask users to verify their identity at sign up. Facebook is one example.

It is therefore difficult to determine whether platforms with more or less anonymity had greater or lesser amounts of abuse, and whether anonymity itself is the driving factor for that abuse. Most platforms allow or enable anonymity – even users with names that appear real could be pseudonymous, and, as we explore later in this report, there are a range of other factors at play, such as the culture of a platform, which are likely to be influencing the amount of abuse.

One of the **expert interviews we did with a policy professional** working in this area described how these additional factors affecting abusive behaviour on a platform make it difficult to isolate the role of anonymity. For example, they felt that the social norms of the platform, which they described as being influenced by the platform's design, stated purpose, rules and other people on the platform, had a significant impact on abusive behaviour.

Level of disposability

In addition to anonymity, the disposability of an account also seemed to have an impact on a user's behaviour, and the two factors often need to be explored in tandem.

The people spoken to during this research often mentioned **highly 'disposable' accounts** – these were sometimes referred to as 'burner' accounts and were used to post abusive content. These are accounts that have only recently been set up and have no pictures, no posts, few followers, and also tended not to be associated with a real name. These accounts were **anonymous** to other users.

In several cases, victims of abuse thought they knew who some of the creators of these disposable accounts were, as they felt they exhibited the same behaviour, said the same things, or even made the same spelling or grammatical mistakes from different accounts, leading them to conclude it was likely to be the same person. So although these users are technically anonymous to others on the platform, sometimes to the victims of their abuse, they are very familiar, if not literally identifiable.

Some respondents also referred to accounts **with established followings and reputations, who posted under a pseudonym.** When being abusive, users with this type of account tended to be more careful not to contravene platforms' guidelines and g banned. More followers meant the account was less disposable, and having the account shut down would be more of a loss to its owner.

Some users with more established followings created temporary, disposable accounts from which they could post abuse without fear of losing their follower base if they were to be banned. Others set up methods to quickly regain followers if they were banned – such as using a specific username or hashtag on their new account to help followers find it, or asking friends to share their new profile to find previous followers.

Section 3
Summary and implications

- The variations and nuances in what abuse and anonymity mean in practice have not been thoroughly explored in previous research.
- There are several types of abusive behaviours and content online which have different characteristics and impacts – as well as being more or less associated with anonymity:
    - Sustained or obsessive – closer to harassment
    - Reactive, angry, one-off
    - Ideology/agenda-driven
    - Intended to discredit or 'call out' individuals
- There are also different degrees of anonymity, which can have very different implications for users:
    - Being completely identifiable
    - Being anonymous to other users but identifiable to the platform
    - Being anonymous to platforms and other users
- In any study about anonymity online, these distinctions must be made and different types of anonymity must not be conflated.

- Some platforms have real-name policies, or require people to verify their details (e.g. using an email address or phone number), but respondents spoke about using false names or contact details to do this. This makes it difficult to determine whether platforms which *appear* to have more or less anonymity have greater or lesser amounts of abuse, and whether anonymity itself is the driving factor for that abuse.
- The disposability of an account seemed to have an impact on a user's behaviour, and should be explored alongside anonymity and abuse. Highly disposable accounts with no followers or other posts were often cited as a source of abusive content, from which users could be abused. If banned, the owner could set up a new account with no significant repercussions or penalties associated with their behaviour.

Section 4
Who benefits from anonymity?

There are a range of reasons why people choose to be anonymous online. This research explored the experiences of people who benefit from anonymity for reasons other than being abusive online, with some feeling they can be online *only* if they are able to be anonymous.

Previous studies show that LGBTQI+[55] people, victims of domestic abuse[56] [57], those with taboo or alternative opinions, beliefs or lifestyles[58], sex workers[59], undocumented migrants[60], teachers, public figures, whistle-blowers and activists[61] may all be considered to benefit from anonymity online.

The types of people who benefited from the protections of anonymity are explored in more detail below.

People who are escaping abusive relationships or families, or want to distance themselves from certain communities

Anonymous accounts were used by people who might face danger or repercussions if their true identities were discovered.

Individuals who had escaped or removed themselves from a difficult or dangerous situation could still be active online under a new or pseudonymous identity – as with Zara, who distanced herself from her family after escaping a forced marriage and did not want them to find her online.

Sally left the strict fundamentalist community she grew up in and used anonymity to join Facebook and other online support forums for advice on chronic conditions she lives with, and to connect with other people who had left the same fundamentalist life as her. Sally felt that without absolute anonymity, she would not have enjoyed the same freedom to engage in these support groups. Her real name coupled with her medical issues and where she is from, would easily identify her to those still in the community.

"There aren't that many people with my name, from my community, with [this health condition] and working in my field. I would have been much slower to engage with [the support groups] if I couldn't be anonymous." **Sally, 30**

Anonymity also allows individuals to hide their online identities from abusive or obsessive ex-partners they do not want to contact them anymore. Sara's ex-partner was obsessive and controlling, and since their breakup, had been stalking Sara online and sending her abusive messages. Sara had struggled to make a profile anonymous enough to ensure her ex did not find it – her ex-partner even tracked down her Etsy page, which had no identifying information other than her artwork.

"I don't want to do the whole fake thing. You know, I just want to be myself and be left in peace and sell some artwork. I mean, obviously, my artwork. I don't put it under my name. It's under different names and things like that. [On] Etsy, she'll find my work. She knows what my work looks like, my style. She's sent me messages." **Sara, 56**

> **Zara was anonymous to hide from her family online**

[55] St James, E., 2020. *Trans Twitter and the beauty of online anonymity*. [online] Vox.

[56] Baker, S., 2016. *Why Online Anonymity is Critical for Women - Women's Media Center.* [online] Womens Media Centre.

[57] SM Kee, J., 2007. *CULTIVATING VIOLENCE THROUGH TECHNOLOGY? Exploring the Connections between Information Communication Technologies (ICT) and Violence Against Women (VAW).* [online] APC WNSP.

[58] Noman, H., 2015. *Arab Religious Skeptics Online: Anonymity, Autonomy, and Discourse in a Hostile Environment.* SSRN Electronic Journal.

[59] Andrews, I., Undated. *A Brief Guide to Online Privace and Security for Sex Workers.* [online]

[60] Stop Online Abuse, 2022. *What is online abuse?* [online]

[61] Van der Merwe, B., 2021. *Are anonymous accounts responsible for most online abuse?.* [online] New Statesman.

Zara escaped a forced marriage in Pakistan, with the help of the Home Office. She had since used a pseudonym on social media and dating apps to avoid being identified by her family and network of family friends.

After the challenges Zara experienced with her family, she deleted her social media accounts and made new ones.

"I deleted everything. I've also deactivated and reactivated since. And I've made everything private. If [I see] a family member or anyone like distant family relative, I just block straight away."

As a single mum who was relatively socially isolated, Zara relied a lot on platforms such as Facebook Marketplace to buy cheaper second-hand goods and various social media groups for support – and felt quite strongly that she would effectively be excluded from much of online life if she were forced to use her real name.

"I changed my name on social media. Even now, I'm on dating apps and I don't have my real name…I don't want to."

People using the online world to explore or develop new lives and identities

Online spaces can be used by people to try new things and express themselves in ways they might feel they are not able to in their offline world.

Some are using social media as a forum where they can be open about their sexuality, gender or ideas, free from judgement by the people they know, or once knew. This includes people in closed or conservative communities, exploring life beyond their current or previous religious or ideological boundaries.

A representative of an organisation supporting victims of abuse shared stories of people from ultra-orthodox Jewish communities, who had relied heavily on online anonymity for this purpose. For some current and previous ultra-orthodox Jews, online anonymity had been used to explore life and information outside of the largely insular, offline world of the community. It had also aided people in leaving the community to begin secular lives – enabling those who leave to maintain distance and anonymity from those they left behind.

Anonymity provides people with a 'risk-free' way to be open about themselves without concern that family or friends would find out. It also provides an opportunity to seek advice and integrate with like-minded people or people going through similar experiences.

Jack, 19, spoke about having very different views from some of his family members, particularly around LGBT rights. For him, anonymity was important so he could engage in conversations around LGBT rights but create distance between those views and the potential for his family to find out about them.

"I might have a certain image with, like, my friends and with my family, [but] the image is just completely different. Being anonymous, I don't have to worry about the image I'm putting forward. I can just kind of speak my mind and not worry about it." **Jack, 19**

Allie set up an anonymous profile while she was transitioning

Allie is a trans woman who set up a Facebook account with a new name before fully 'socially' transitioning.

"I can remember that I just I went on Facebook and looked for like transgender support groups and had just joined a couple… Initially, when I did that, that was on my regular Facebook account. And then after like a week of doing that, I was thinking as I was getting ready to come out to like, my family; 'What if they've already seen that?'. So, I just stopped using that account and I replaced my picture on it with an album cover and made all my old pictures private. At the same time, I created a new account… It was semi-anonymous, I guess."

Allie felt that being able to have two Facebook accounts while transitioning was critical to her transition, allowing her to come out and find online support groups without people who did not know she was trans finding out.

"I got a lot of support through Facebook, it was completely essential to me at the time. Because if I would have been left to my own devices, still just to talk with my family – that's where all of the problematic conversations are coming from – and if I did not have any other way to link in with the trans people, because it was all online, it would have felt immensely isolating."

She also made meaningful connections with other anonymous users on dating apps who reached out to her for help and advice.

"I've made friends with a few trans women on [Grindr]. There's this girl who is not out to her family, we've become quite good mates. But like, she's not the only trans woman that I talked to like that on Grindr. There is like two or three others."

Allie thought anonymity was a really important tool for people who were coming out and need to find support or explore that part of themselves safely. She felt that without anonymity she would not have been able to access the support online that she needed in the early days of her coming out.

She had concerns about ID verification as this would make it very difficult for trans people who may not have ID in their new name but wanted to be able to post on a platform.

People who engage in debate and discussion on sometimes contentious topics online

Anonymity gives some people the confidence to express their opinions or views in situations where they might otherwise feel embarrassed or worried that this might have negative repercussions.

Some people we spoke to were using anonymous accounts to engage in debate or discussions that they either did not want other people to see them doing, or they felt more confident doing so under a pseudonym.

Jack got into heated debates online on topics to do with race and LGBT issues – he did not feel confident doing this under his real identity as he felt people might focus on his ethnicity and appearance. Jack is mixed race but says himself that he does not "look very black" and thought people might focus on aspects of his appearance and assumed identity when they were discussing contentious topics. He felt this would become a point of focus in debates about racism. *"I feel like being anonymous is better because, like, when I make comments about race, despite being black, I just don't look very black. There is often this sense of 'You're not quite black, though.' But if I say it anonymously, it's fine because, well, I actually am black. So that judgement just is not there."* **Jack, 19**

Jack also has radically different views from some of his family, and valued being anonymous on certain platforms so as not to be identifiable to them. He had strong views about LGBT issues, but said his family on his dad's side were homophobic. He liked being able to have some control over who saw things he posted about. On Instagram, he reposted things on his story and used the 'close friends' feature, so his dad's side of the family did not see.

Similarly, Ibrahim enjoyed engaging in heated debates and he used to use anonymous accounts because he did not want to be "seen" online. He had been bullied a lot when he first moved to the UK from Dubai as a teenager and did not want this to happen again during his interactions online. He also did not want his strong views to be seen by colleagues or employers as he worried this could affect his career.

"I just used [anonymous accounts] because I did not feel confident putting myself out there as a person to be made fun of… If you put a picture up someone will use it in a negative way…" **Ibrahim, 22**

Darren, in his late thirties, has an anonymous profile on the platform Gab. Although he sometimes posted abusive content online, he had other motivations to be anonymous. He was a strong proponent of free speech and felt he and others could be "persecuted" for their views if they were not able to be anonymous.

People who want to keep their work and personal lives separate

Some people work in industries or professions that might make it dangerous, uncomfortable, or problematic for them to use their real identities online.

Examples of this could be journalists working on controversial or dangerous topics; sex workers who want to keep their work separate from their personal lives, either to keep their profession private or to ensure clients cannot discover personal information about them; or teachers who want to ensure their pupils or employers cannot find them online.

"I comment and post on some of the teaching groups I'm part of. I wouldn't do that, I'd definitely not do some of that, like what I say on there, if I did not have anonymity. I was on sick leave at one point, whatever questions I had about that I would have been entirely uncomfortable posting because I did not want anyone to see it, did not want my school to see it." **Sally, 30**

Cat did not want her career to be associated with her real identity

Cat is an adult content creator. She works hard to maintain a clear divide between her work identity and her personal identity, and any threat to that is extremely concerning to her.

A couple of people in the past have been able to work out who she "really" is and threatened to contact her family.

"I get a lot of people sending my mum my content…Somebody got my mum's number and pretended to be me… They want to ruin my life, and I'm like: 'You're not going to ruin my life.'"

Cat uses a pseudonym across all of her social media and work accounts to create a public-facing identity. She said having to use a real name would be "the end of it all" and she would not be able to do her job.

Anonymity also provided Cat with the security of disposability, the ability to later move on from this part of her life and pursue other things. For Cat, her adult content creation is a means to an end right now, and she expressed wanting to pursue other endeavours in the future. Cat felt that having her real name attached to adult content creation would make this difficult.

"The reason I use an alias is because when I want to get rid of that part of my life, I can just do it. Whereas otherwise people would know me as '[real name] the porn star' for the rest of my life."

Some people who benefited from anonymity also received anonymous abuse

Many of the respondents who were benefiting from the protections of anonymity had, at different points, received abuse from other anonymous accounts.

Jay, 21, had used an anonymous account online, and while using this account would get into heated discussions with people he thought were being racist or offensive. While using this anonymous account he would receive abuse from other anonymous users.

"Sometimes it would get taken to the DMs. I remember this one guy was just hurling abuse…this guy says black people shouldn't exist…Anything they could think of that they felt might offend someone, they would send that my way." **Jay, 21**

Nonetheless, these people maintained that anonymity online was a precious and important tool that they highly valued. They felt that any move to reduce or ban anonymity would have big negative implications – if not for them, for other people who required anonymity to be safe online.

"I feel like we shouldn't take that away… we should allow people to exist as nobodies… where there could be repercussions for being politically active, regardless of whether they're being negative or positive." **Jack, 19**

Few expressed concerns over verifying their identity to a platform or third party

We presented interviewees who were anonymous to other users – using a pseudonym and usually with an unidentifiable profile picture – with three different degrees of anonymity and asked them which they would be comfortable with:

A) Being anonymous to the platform and other users
B) Being anonymous to other users but identifiable to the platform / a nominated third party
C) Being identifiable to everyone

While all of them wanted to remain anonymous to other users for the reasons outlined above, few expressed any concern about being identifiable to the platform itself, or to a third party.

Commonly, this group of users were concerned with shielding their identity from other users, not the platform itself, and were not opposed to measures designed to hold people to account for their behaviour online.

However, having to identify yourself to the platform might be challenging for people from marginalised communities. For example, for transgender people, their 'legal' name may no longer be the name they use. Allie, a transgender victim of abuse and someone who had been anonymous online, felt she would have struggled to access online spaces if she had needed ID.

"Without anonymity I don't think I'd have been able to get the support I needed when I was coming out. If people had to use ID to verify themselves I think it would be very tricky for trans people who may not have ID in their name."
Allie, 33

The people who did express concern about being identifiable to a platform included some perpetrators of abuse. They were worried platforms could use this information to censor free speech. Darren, who mainly used Gab to post inflammatory content, cared a lot about maintaining freedom of speech and he felt anonymity was very important for this. He was strongly against any form of verification or handover of personal data. Darren believed that the government would "abuse their powers" if they enforced user verification and platforms had to collect more information on their users - he felt that the government would use the information users shared with platforms to arrest any users saying things they did not agree with.

"Not a fan - I already keep my speech on here within the lines for the most part. I'm not an idiot; if they want me, they can get me, and I'm not giving them a reason by throwing around "hate speech". But anytime the government gets powers like that, they [the government's powers] are abused [by the government], no exceptions." **Darren, 35-40**

Section 4
Summary and implications

● There are a range of reasons someone might want to be anonymous online, which aren't linked to enabling abusive behaviours. We spoke to a number of people who were using anonymity for their own protection or to feel more confident in expressing their true selves online.
● Removing or reducing the ability to be anonymous on social media platforms would therefore likely have a huge negative impact on individuals in these situations. An inability to be anonymous on a platform could mean that vulnerable individuals are excluded from parts of the internet.
● However, people who benefited from anonymity were largely benefiting from being anonymous to other users and most would accept being identifiable to the platform itself.

Section 5
What role does anonymity play in abuse for victims?

Victims had received abuse from accounts with varying degrees of anonymity

During interviews, victims showed researchers screenshots, photos and live examples of the abuse they had received.

While it is not possible to claim this is a representative sample of online abuse, it highlighted the range of accounts that abuse was coming from. These accounts ranged from being fully identifiable (to other users and likely to the platform), to those which were wholly anonymous to everyone.

Typically, abuse came from accounts that fell into three categories, detailed below.

Much of the abuse victims received came from highly disposable 'burner' accounts

A lot of the abuse that people received came from highly 'disposable' or so-called 'burner' accounts. These accounts typically appeared to be recently set up and had no pictures, no posts, few followers, and no identifiable information. On some platforms it was possible to verify that the account was recently set up, as information about when a user set up their account was shown in their user facing profile information. However, on other platforms this information was not visible to other users, but respondents often assumed the accounts were recently set up based on a limited amount of content on their profile and few followers or accounts they were following.

The reasons behind perpetrators choosing to use these 'burner' accounts are explored in more detail in section 6. To summarise, perpetrators were often posting abusive content from disposable accounts, which were separate from their "real", more permanent accounts. This meant that if the abusive content violated platform guidelines and the account got suspended, there were no consequences for the perpetrator's "real" account, and they were able to remain on the platform using their "real" account (or using a new, disposable account).

Some victims of abuse believed some perpetrators had chosen to create a number of disposable accounts to target victims with a high volume of abuse from seemingly separate users. However, none of the perpetrators we spoke to as part of this research engaged in this specific behaviour.

The characteristics of these disposable 'burner' accounts include:

● **Anonymity to other users:** Most likely to be unidentifiable to other users, using an unidentifiable user name and photo.
● **Anonymity to platform:** Likely to use an unidentifiable 'burner' email to set up account, and not give any real information to the platform
● **Disposability vs stability of account:** Disposable, unlikely to be used for a long period of time (i.e. may be used for as little time as a day or a few hours)

Fiona receives significant amounts of online abuse from what they thought were anonymous 'burner' accounts
Fiona is shaping their career and much of their online identity around their passion for LGBTQ rights and social justice. They are also an advocate for body-positivity and self-described as a "fat activist".

As a teenager, Fiona experienced a lot of serious bullying at school which greatly damaged their confidence and mental health. Fiona has struggled severely with anxiety and depression in the past and has been open about this on social media, including about a suicide attempt.

They are the victim of a large amount of abuse on TikTok and Instagram. Fiona spoke about receiving abuse from anonymous 'burner accounts', set up specifically to send abusive content.

"It takes about three minutes to set up a new account when you're blocked. And some of these trolls are so motivated. They're like bees to honey."

While 'burner' accounts were typically anonymous to other users, some victims believed accounts belonged to the same user

A few victims thought they recognised patterns in behaviour or language used in the abuse they received, such as common misspellings of words or common use of certain words. This meant they believed some anonymous and abusive accounts belonged to the same user, even if they did not know their identity.

"I used to call him my alphabet stalker… he would always pop up with a brand new account, say some hateful rubbish and the mods would block him… it was always the same phrasing, always the same way of speaking, always the same spelling errors." **Angela, 47**

"You block one, and then they come back, either under a new comment or in your personal inbox, like 'You think you can get rid of me?' New account, same conversation. You can see the language they use, the tone or the spelling." **Jodie, 24**

Some received abuse from established accounts using pseudonyms

Some received abuse from accounts with an established reputation or following, but who posted under a pseudonym.

The characteristics of these types of accounts include:

- **Anonymity to other users:** True identity unidentifiable to most users, although some users may be aware of who is behind the pseudonym
- **Anonymity to platform:** Some accounts may use their true identity when signing up to the platform, while others may not
- **Disposability vs stability of account:** Stable, likely to be used over a period of time and have built up a following

Dana received abuse from someone who was well-known for being rude and cruel online
During the pandemic, Dana's younger sister got into an argument online with someone from her local area because she felt that they were sharing content about Covid regulations that was racist toward Chinese people. Dana stepped in and got sent abusive messages from the abuser and a group of his friends. The pile-on lasted for a few days and prompted Dana to delete her Twitter account.

She explained that the perpetrator had a well-known pseudonymous account which he used to post "edgy" memes and content that was offensive or borderline offensive. While Dana and others who knew who this person in real life were aware that this pseudonymous account belonged to him, other users who did not know him would not necessarily be able to identify the person behind the account.

He had 'pinned' one of the tweets in which he mocked Dana to the top of his profile: Dana felt that he was proud of his account's reputation for "edgy, offensive" content.

Victims also received abuse from non-anonymous accounts

A proportion of the abuse that people received came from non-anonymous, identifiable accounts.

The characteristics of non-anonymous accounts include:

- **Anonymity to other users:** Identifiable
- **Anonymity to platform:** Real identity used to sign up to platform
- **Disposability vs stability of account:** Stable, likely to be used for a prolonged period of time

"You get really obvious [messages], people just send me a picture of their penis, very explicit sexual comments and messages like that. It irritates me and it angers me that we have to deal with those sorts of things, particularly as women…Interestingly, most people who send me [male genitalia] pics and explicit messages that are sexual or harassing in that way aren't anonymous." **Jodie, 24**

Sara knew her online abuser

Sara's long-term relationship with her controlling and abusive partner ended a number of years ago. Since, Sara has received intermittent online abuse from her ex-partner via several channels including email, Facebook, Etsy and the Nextdoor app.

Sara's partner is identifiable and sends the abuse using her own full name, stalking Sara online seemingly in an attempt to uncover any new accounts she may have opened.

Sara has deleted her accounts across Facebook and Nextdoor, and is not engaging whatsoever in any other online platforms. Sara previously generated income by selling her artwork online, but due to the abuse she received from her ex, she can no longer do this. Although Sara used a pseudonym, her ex-partner recognised the artwork and Sara believed she would continue the abuse and harassment even on platforms such as Etsy or eBay.

Abuse from 'anonymous' accounts affected victims differently depending on the type of abuse

Some of the victims were able to describe how abuse had affected them differently depending on the degree of anonymity associated with the account it was coming from.

For some victims, whether anonymity increased or decreased the impact of abuse differed according to the *type* of abuse they had received.

Reactive, angry abuse was taken *less seriously* by some when it was from anonymous 'burner' accounts

Reactive, angry abuse was taken less seriously by some victims when it was from anonymous 'burner' accounts, compared with when it was from identifiable or pseudonymous accounts. Some assumed that when this type of abuse was coming from anonymous accounts it was from a "bot" or "foreign account", which meant they took it less seriously, describing it as feeling "less real", and finding it easier to ignore.

"It could be anyone so you cannot take it that seriously – it could be a bot… you can just block the account and move on" **Dana, 28**

"I wouldn't even respond. I'd just block them." **Angela, 47**

"It's easier to sweep it under the rug and forget about it when it's from an anonymous account." **Jodie, 24**

Emily described feeling more worried when abuse was coming from an account of a higher profile; one that was more public, more identifiable and had more followers, compared to an anonymous burner account with few or no followers. The risk of a potential pile-on or further abuse or shaming was perceived to be greater, as other people might see it.

However, Bill felt more able to rationalise abuse from identifiable accounts:
"If there's an account and you can see them – they've got a name, you can see the kind of things they post – you can say 'OK, maybe they did not have the kind of education where these things were discussed around them or they're from a background where this sort of behaviour is more regular'… with an anonymous account I cannot do any of that… I cannot explain it to myself… it makes it stronger because it grips hold of you… it's almost like the anonymous accounts are playing mind games with you." **Bill, 27**

For some people, sustained, obsessive abuse was felt to be *more intimidating* when it was from anonymous 'burner' accounts

On the whole, sustained, obsessive abuse was perceived by some victims as *more* intimidating and threatening when it was from pseudonymous or anonymous accounts, as opposed to identifiable accounts. This was due to the threat of the unknown, and the perception that they cannot stop it as abusers will keep creating new accounts.

"You have no idea who it could be, and you know there is nothing you can do to stop it." **Fiona, 25**

One organisation which supports victims of abuse spoke about the people it supports feeling "reassured" and "relieved" when victims know there is something that can identify their abusers (such as an identifiable email address), as they feel there is more of a chance the behaviour can be "addressed".

What did victims think about the ability to be anonymous online?

Victims expressed a range of opinions about the ability for people to be anonymous online, often weighing up the role it played in abuse they had received with the benefits that it had offered to them and others.

Some identified anonymity as a driver of abuse

Some victims spoke about the role anonymity played in the abuse they had received – people felt it made it "easier" for others to be abusive, particularly referring to 'respawner' and or 'burner' accounts. When respondents spoke about 'respawner' accounts they were referring to accounts set up by a user who had previously been suspended or banned from a platform, and then reappeared on a new account.

"I think [reducing anonymity] would moderate people's behaviour a lot… I don't think people would say half the stuff they say online without it." **Angela, 47**

A few people separated the ability to be 'anonymous' from the ability to create 'burner' accounts, and felt that reducing the ability to create 'burner' accounts would have a bigger impact on reducing abuse, while protecting the benefits of anonymity.

"Anonymity is important, especially for minority groups… it shouldn't be so easy to make burner accounts though." **Fiona, 25**

Many victims had benefited from the ability to be anonymous online and had concerns about restricting it

As identified above, many of those who were victims of anonymous abuse, had also benefited from the ability to be anonymous online, or were able to recognise the benefits that it had for others, especially for minority groups.

The majority of victims felt that the ability to remain *anonymous to other users* online was particularly important.

"There are lots of people who really need it [anonymity] and it should be protected – like sex workers or domestic abuse victims. Verifying to the platform, but being able to be user facing anonymous would probably be okay." **Emily, 30**

While the ability to remain anonymous to other users was seen as important to most victims, most felt comfortable with the platform knowing their true identity. However, there were a few victims who had

concerns about a perceived degree of "power" that it could give to platforms, and the difficulty for some people to verify their identities.

"I would want to see a separate body in charge of that… not anything to do with Facebook… they have got a history… I wouldn't trust them." **Bill, 27**

"I would feel comfortable giving ID… I'd feel safer… but if someone does not have ID… it's indirect discrimination… there are social issues here at play." **Bill, 27**

A few thought that restrictions in anonymity would prevent them receiving some of the abuse they did, but felt conflicted, because they recognised that the costs of this often outweighed the benefits.

"On the flip side… as a trans person coming online is often the first place that you say anything and the first place you get to explore and a lot of people do that on anonymous accounts… so in that respect anonymity is good." **Angela, 47**

"Anonymous accounts can be good…there are people out there who use them for good, like exposing abuse or sharing stories like that. For particular reasons like that it protects people who wouldn't have a voice otherwise." **Emily, 30**

How did victims manage and prevent abuse?

While identifying victims and perpetrators of abuse to take part in this research, we searched for abusive content online. Abuse was not easy to find, particularly on 'mainstream' social media platforms. However, this was not a systematic search or measurement study, as finding online abuse was not the main purpose of the research, but rather an observation.

Additionally, many perpetrators and victims of abuse spoke about and showed us screenshots of abusive content that had since been deleted by social media platforms (the abuse victims showed us tended to occur on the more 'mainstream', larger social media platforms). This suggests some platforms are effectively locating and removing large amounts of abusive content online.

However, people still received and viewed abusive content online, some in significant quantities. Efforts were required on the part of victims to manage and prevent themselves from seeing this abuse, with some feeling the onus was unfairly on them to protect themselves.

Many had blocked users to prevent abuse, but this process could feel futile

Victims spoke about various strategies they had used to manage and prevent online abuse. Many had blocked users who were abusive, or deleted abusive comments on their profiles.

"I will try to remove any [abusive] comments as quickly as possible, report it and block the account. I don't want people who follow me especially young people to see it" **Jodie, 24**

Angela had made friends with other TikTok users and they would assign one another as moderators on their TikTok live streams. When Angela received abuse from anonymous burner accounts, her moderators would immediately block each one.

However, Angela and other victims spoke about the challenges of blocking users who would then reappear with new 'burner' accounts and continue the abuse. This meant the process of blocking users could feel futile and time consuming.

One user had used a third-party 'bulk blocking' tool to try and overcome this. The online tool enabled them to automatically block any users related to an account they had blocked (i.e. automatically blocking any users who followed the account or were followed by the account).

We spoke to some people who had public profiles and a large number of followers (e.g. 10k plus followers). These people received huge volumes of messages and comments, so felt that tools such as being able to block users and remove content were ineffective and labour intensive.

Some wanted to save abuse as 'proof' or evidence

Although many blocked and removed abusive content, a few also wanted to capture the abuse as proof or evidence, and would take screenshots of comments or direct messages they had received. A few people would then share these images with their followers, to raise awareness of their abuse, which often elicited positive responses, sympathy and support (this is discussed in detail in section 7). However, others deleted abusive content and did not want to have a record of it because it was upsetting.

Victims also spoke about how abusive content had been removed by the platform, and were often unable to find abusive content they had previously seen. Although this was positive, even when content existed for a short while it could still be seen by the victim and potentially do harm.

Some were reluctant to use tools which may prevent them from seeing abuse if the abuse was still visible to others

A few were aware of tools which enabled users to hide comments or posts containing certain words or phrases that they did not want to see. However, if this content was still visible to others, some felt anxious that users might be talking about them in content which was hidden, and would be unaware of what was being said about them. As a result, some said they would rather know what was being said, even if it contained words or phrases they did not want to see, than choose to hide it.

Similarly, Dana, described feeling "paranoid" once she had blocked her abusers on Twitter, as she was not sure what they were saying about her now she could not see their posts.

"Even once I'd blocked the [15] accounts I still felt paranoid about what other people were seeing and what they were saying about me." **Dana, 28**

This was particularly the case for those with higher profile accounts, who felt that these tools may limit their ability to engage with others in the way they wanted to.

Some experts raised the difference between traceability and anonymity when it came to managing abuse

A few organisations (those who support victims of abuse, as well as think tanks, research and campaigning organisations), raised the difference between anonymity and traceability – as a way to distinguish between users being identifiable to other users and the platform, and being traceable by law enforcement.

Many felt that traceability was key when it came to dealing with abuse.

"We generally don't care about people being anonymous to other users as long as they are not ultimately anonymous … if they engage in criminal conduct they lose their right to anonymity and should be able to be tracked down [by law enforcement]" **Organisation supporting victims of abuse**

Another research and campaigning organisation described anonymous abuse as more "frightening" for victims and "harder to resolve" due to the lack of traceability.

Indeed, organisations supporting victims of abuse raised the challenge of knowing who is responsible for posting certain bits of content online, even where platforms were able to locate individuals through data they hold, such as the IP address of the user[62].

Section 5
Summary and implications

● Given that victims receive abuse from a range of accounts, with varying degrees of anonymity, restricting the ability to be anonymous online may prevent some, but not all, of the abuse that people receive.

---

[62] Internet Protocol address

- However, it is important to separate *anonymity* from *disposability*. The 'anonymous' abuse victims received was often from highly disposable 'burner' accounts – which were recently set up, with few posts or followers, and did not appear to have the intention of being used on a long-term basis. Making it harder to set up these disposable 'burner' accounts, is likely to significantly reduce the amount of abuse victims receive.
- Many victims had personally benefited from the ability to remain anonymous online, or recognised the benefits for others. Protecting the ability to remain anonymous to others, as opposed to being anonymous to the platform, was seen as particularly important. Future solutions must find a way to balance protecting the ability to remain anonymous online, while also reducing abusive accounts. For example, reducing the ability to quickly set up new accounts, as mentioned above, would likely decrease abuse without restricting the ability to be anonymous.

Section 6
What role does anonymity play in abuse for perpetrators?

Although people were using anonymity for a range of reasons, some people who had anonymous accounts were being abusive online.

Some were using anonymous accounts to abuse individuals without real-world repercussions

The 'disinhibition effect'

There were examples of the 'disinhibition effect' – with a few people feeling they were only able to be abusive because of the option of remaining anonymous to other internet users online. They felt less accountable, and freer to say whatever they wanted.

For example, Jack was concerned that if he was not anonymous on Twitter, a future employer may see some of the "contentious debates" linked to his account, which he felt could affect his employment. Jack reflected that the ability to be anonymous enabled him to be more "rude" and "confrontational" online.

"I'd probably not be so aggressive if I was worried – well if it was attached to my name. It's like, you know, what if a future employer sees this…that kind of stuff. Being anonymous can bring out the worst in a person" **Jack, 19**

Alan also took part in heated arguments online, and although not anonymous himself, reflected on the impact anonymity had on other people's behaviour during those arguments.

"Because of the anonymity afforded to a lot of people, that makes it very easy for them to take extreme positions and be more abusive, or more in people's faces than they would if they were not anonymous" **Alan, 60**

> John Suler's theory of the disinhibition effect, suggests that online communication causes users to lose inhibition and act in ways that they would not offline. This can work in two ways: benign disinhibition (unusual openness, vulnerability, generosity) and toxic disinhibition (rude language, hatred, harsh criticisms).[63] The principle is that when a person's online persona is not traceable back to their offline self, people feel freer to act differently.

Some people who were abusive did not want to be anonymous

Not all of the abuse seen in this project was coming from anonymous (by any definition) accounts. And sometimes people were actively keen to use their real names as they felt there were benefits from being abusive and identifiable.

Two of the people we spoke to who acted in abusive ways online, Lorraine and Sia, were particularly keen to remain identifiable. They often felt their behaviours were legitimate or brought them valuable benefits, and they did not want to hide this.

---

[63] Suler, J. 2004. The Online Disinhibition Effect. *CyberPsychology & Behaviour* 7(3) 321-326.

The abuse brings status

Sia, 25, felt their actions online brought them status: their inflammatory persona online was an important part of their identity.

Sia gives abuse to other users in political discussions on Facebook and Twitter. For example, they sent abusive messages to users in debates about the position of asexual people within the LGBTQ umbrella.

They created anonymous 'burner' accounts to enter private Facebook groups they had previously been banned from, where they would screenshot posts by users who they thought were "stupid". They would then share these screenshots from their non-anonymous account in other Facebook groups amongst peers.

This behaviour would give Sia status amongst those peers: status came from "taking down" alternative views. They described the "unique rush" they get "when people say something silly and you can make fun of them".

Sia only used anonymous accounts as a practical step to evade bans from Facebook groups. Aside from this, they had no desire to be anonymous: their abusive online behaviour was rewarded by peers and was an important aspect of their identity. Sia was proud of their opinions and believed they were in the right. They wanted their profile, opinions and abuse to be attached to their real identity.

"it's more fun to argue with people when you know they're real…can I really be bothered to argue with some no picture, two followers guy…no one really cares what he says and I don't either" **Sia, 25**

Lorraine, 56, also feels a sense of status and recognition for what she says online, and it was important to her that she used her real name. Lorraine uses social media to express her opinions relating to immigration and people from different religions and ethnicities. She wants to meet people with similar views and has made friends in this way.

She also used to tweet presenters from LBC and Talk Radio, who would occasionally read some of her Tweets out on air (though not the most strong expressions of her opinions or abusive content) relating to her views on immigration. She describes the feeling of recognition when someone would read these tweets and say her name. Now that she has had to stop using her real name (due to an account suspension), she feels frustrated that the presenters do not know it is her tweeting in anymore.

"That annoyed me cos I had to change my name when they were reading them [my tweets] out, but they don't know who it is, they don't know it is the same person they were reading out last year, which absolutely guts me that they even they don't know my real name now. He used to always know who I was, but I've lost all that as well now" **Lorraine, 56**

The abuse is a way to share "important truths"

Both Lorraine and Sia felt they were in the right when commenting online – that their opinions were correct and truthful. Because of this, they did not want to "hide" behind an anonymous account.

Similarly, James, 72, used websites like Twitter, Gab and Gettr both to "seek out" and post about "truths", particularly around vaccines, global politics, conspiracies and evolution. James was not interested in being anonymous.

During the process of trying to find people to take part in this research, we observed many users on alternative social media sites, such as Gettr and Gab, who had been banned or suspended from Twitter. Many of these users were frustrated because they felt that their views had been "censored". For some, they felt their suspension acted as "proof" that platforms were obstructing free speech.

For example, one individual on Gab referred to Facebook as "facistbook" and spoke of a "purge of right-wing views" on Twitter.

Anonymity was often an outcome of abusive behaviour, not a precursor

Creating an anonymous account online (to other users and/or the platform) often appeared to be a consequence of abusive behaviour, rather than a driving causal factor. Being abusive on many platforms can result in account suspension. Perpetrators of abuse were often creating anonymous accounts after they had

been abusive on their identifiable account and been banned, as they could no longer access the platform with their real details.

Some had created anonymous accounts after their identifiable accounts were suspended or banned

Some of the perpetrators in this sample had been suspended or banned from using their non-anonymous profiles, after repeated violations of a platform's community guidelines (primarily on Twitter and Facebook). They described creating new, anonymous profiles in order to remain on the platform. In these instances, perpetrators were more concerned about being anonymous to the platform, to avoid being caught for evading suspensions, rather than to other users.

Similarly, people like Sia were using anonymous accounts specifically to gain access to groups they had been banned from.

Think tanks and research organisations interviewed for this research also spoke about the phenomenon of people creating anonymous accounts as a response to being banned from a platform, or in order to avoid being banned from their main account.

One expert from a think tank and research organisation had a team of volunteers tracking abusive accounts on Twitter, to identify whether the accounts were likely to be owned by the same person. For example, by looking at specific types of language or content they were posting, and who they were following, they identified whether the same person had 'respawned' on a new account. They identified some users with up to 40 accounts that had been suspended.

Fred's journey to becoming anonymous online

Fred is in his mid-20s and lives in the US. He is currently a student but also works as a security guard.

Fred is an example of someone who became anonymous because he wanted to stay on the platform he was using and could no longer do so being identifiable.

Fred uses Twitter frequently. Until around two years ago, he was fully identifiable to the platform and to other users. He had his full name on his profile. He primarily enjoys "baiting" people online who he considers overly sensitive – or people who are contributing to – in his opinion – an "overly censorious" discourse.

Around two years ago, he tweeted "All unvaccinated people should be thrown into camps". He described this as a joke but was suspended from Twitter as a result.

Fred loves being on Twitter, and was keen to stay on the platform, so he quickly set up a new account under a fake name – "I cannot keep respawning with my real name".

To create a Twitter account an email address is required – Fred described the process of easily setting up a new email address, which meant the platform could not link him to his previous profile.

Fred's behaviour and attitudes were the same when he was identifiable as when he was anonymous. The ability to be anonymous did not seem to drive his behaviour. Instead, becoming anonymous was a means to remain on a platform he had previously been banned from and continue his abusive behaviour.

Lorraine's journey to becoming anonymous online

Lorraine is 56 and lives with her two dogs. She is not currently working and is receiving Universal Credit, but previously worked in admin.

Lorraine spends most of her time online – finding people who have similar hobbies and interests to her. For example, people who talk about TV programmes she enjoys. She also spends a lot of time talking about her views on immigration. Lorraine expressed frustration that "no one is doing anything about immigration" and that anyone who shares anti-immigration views is branded a "racist".

Lorraine has had three Twitter accounts. She was initially identifiable to other users, but ended up being anonymous to users *and* the platform after multiple suspensions and bans. Lorraine does not want to be anonymous on Twitter, but felt she had to use fake details in order to evade immediate suspension again.
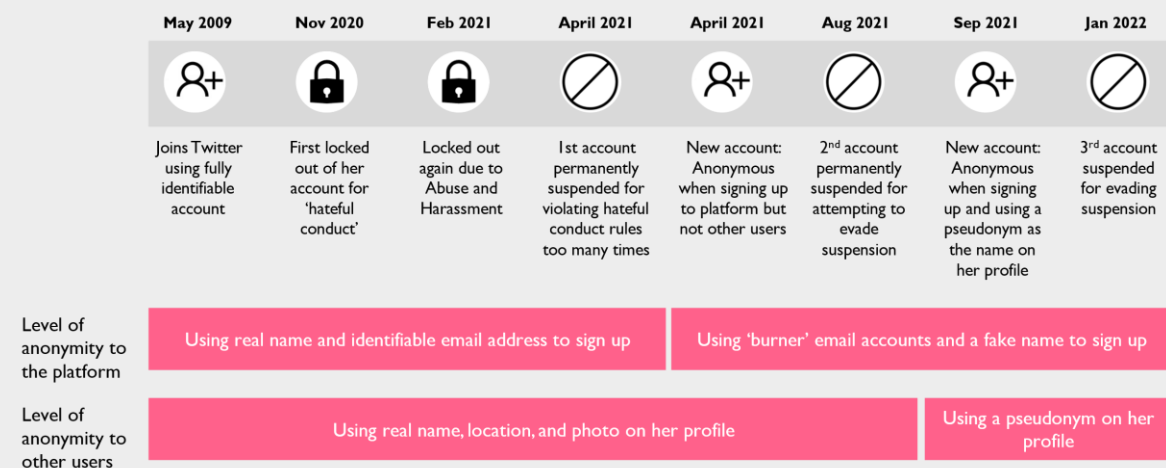
May 2009	Nov 2020	Feb 2021	April 2021	April 2021	Aug 2021	Sep 2021	Jan 2022
Joins Twitter using fully identifiable account	First locked out of her account for 'hateful conduct'	Locked out again due to Abuse and Harassment	1st account permanently suspended for violating hateful conduct rules too many times	New account: Anonymous when signing up to platform but not other users	2nd account permanently suspended for attempting to evade suspension	New account: Anonymous when signing up and using a pseudonym as the name on her profile	3rd account suspended for evading suspension

Level of anonymity to the platform

Using real name and identifiable email address to sign up | Using 'burner' email accounts and a fake name to sign up

Level of anonymity to other users

Using real name, location, and photo on her profile | Using a pseudonym on her profile

Figure 5: Lorraine's journey to becoming anonymous online

Others had pre-emptively set up anonymous accounts to avoid suspensions

Others had pre-emptively set up separate, anonymous accounts, in addition to their 'main account' on a platform. This enabled people to avoid having their main accounts suspended for violating the platform's guidelines. They seemed to view their 'main' account, which would likely have followers or friends and records of previous posts, as indisposable, compared to their 'burner' anonymous account, which was seen as disposable.

For example, Fiona received the abuse [*Figure 6*] from someone who admitted to creating a separate account to abuse them on, to avoid their main account being banned.

Replying to

I have created a separate account to say this so I don't get banned on the other one. But you are so ████ UGLY. It makes my stomach turn. You look like the ████ grinch. You are an awful person

20:04 · 18/11/2020 · Twitter for iPhone

Figure 6: A tweet sent to Fiona demonstrating disposable burner accounts

The ability to set up new, anonymous accounts so easily may be creating a different type of 'disinhibition effect'

This suggests that some perpetrators' behaviour is 'disinhibited' as they do not worry about breaking platform guidelines.

We heard how easy it was to set up new accounts when perpetrators were suspended, which seemed to create a different type of 'disinhibition effect' – allowing people to behave abusively without as much concern that their account would be suspended, as they could quickly set up a new one.

"You can just set up new Gmail accounts which then generates new numbers [to use to create a new Twitter account]. So you can basically have as many as you want. Once you know how to do it, it's easy." **Fred, early 20's**

As described in Section 4, many respondents could see problems with restricting the ability to be anonymous, but few felt there were downsides to restricting the ability for users to quickly create multiple new accounts.

Section 6
Summary and implications

• There were reasons for some abusers to actively want to remain identifiable. For example, the status the abuse brings them, or their belief they are posting "truth" and not abuse.

• Anonymity was sometimes a consequence of abuse, rather than a driver. In this case, perpetrators had set up anonymous accounts after being suspended from platforms due to abusive behaviour on their identifiable accounts.

• It is important to account for all drivers of abuse, recognising the role that platform design and other motivators play – which may have a more significant impact than anonymity on abusive behaviours in some instances – in order to effectively reduce abuse.

• The ability to quickly set up a new, anonymous, account may be creating a different type of 'disinhibition effect' – allowing people to abuse others without fear their account will be suspended or sanctioned, as they are using an account which is effectively disposable.

• This suggests that more work is needed to ensure those who are abusive, or have been suspended from a platform, are not able to create new, anonymous accounts. This is highly likely to prevent anonymous abuse, while protecting the benefits of anonymity.

Section 7
What other factors play a role in shaping abusive behaviours?

Anonymity was one of many factors which were seen to correlate with abusive behaviours. However, there were other factors supporting or enabling abusive behaviours on social media – some of which respondents felt should be acted upon to reduce online abuse. This includes the **disposability of accounts enabling repeat offender abuse** (see pages 34 and 42 specifically), which has been discussed throughout the report.

Some of the other factors identified throughout the research, which play a role in shaping abusive behaviours, are explored below.

Reward mechanisms intrinsic to social media platforms may also perpetuate cycles of abuse

Platform design itself is well known to impact user behaviour[64]. Reward mechanisms, such as likes and follower counts, incentivise users to do certain things and may be playing a role in incentivising abusive cycles of behaviour.

Some of the perpetrators of abuse were incentivised to post shocking and inflammatory content to get attention, and being rewarded with 'like', 'view', 'comment' and 'retweet' mechanisms:

"That was my best ever tweet I got 135,000 views, moaning about [prominent footballer] […] that was my best ever reaction, that went up sky high every day the amount of people who saw it and retweeted it" **Lorraine, 56**

"I developed this tendency to try to want to see people get angry because it made me feel like I was winning… Even in the context of Twitter, if you successfully ratio someone [flooding a post with abuse, disagreement or distraction that outweighs any support or likes] that's a win… It's funny and we laugh at them together…" **Fred, 20s**

A few perpetrators found that content which "pushed the boundaries" or was likely to annoy other people, would likely generate a reaction, giving them attention and raising the profile of the content.

"It was rude but I thought it was great content because a lot of people in my life knew who he was… so it was fun" **Sia, 25**

This was true even if some of the reaction was negative such as other users expressing their outrage or disagreement with the content. Fred explained that it was especially important to post content that "pushes the boundaries" when trying to regain followers after being banned:

"The only way to get your Tweets out on the timeline is if they get engagement" **Fred, 20s**

The consideration of these wider platform design mechanisms and the impact they have on abusive behaviour was also raised by experts. One think tank and research organisation felt that safety by design (designing a platform to reduce the risk of harm to those who use it) was the 'only solution' to reducing abuse. They felt that anonymity should be considered as one factor in a much wider safety by design process.

Some people's online identities are associated with their status as victims of abuse

A few of the victims participating in the research regularly shared the abuse they received on their profiles in order to 'call out' the abusers and fight back against it. For example, sharing screenshots of abusive comments they had received on a recent post, or direct messages they had been sent.

[64] GOV.UK. 2021. *Principles of safer online platform design.* [online]

Often, these types of posts meant the victim received an outpouring of support and attention. For some of the research participants, much of the content they shared online related to their status as a victim. A few had even built their careers around their identity as a person who receives a lot of abuse.

While victims of abuse should be able to react to it in any way they see fit, the sharing of the abuse they receive, and subsequent positive reaction of their followers, does suggest there may be an incentive to receive abuse.

While this could in fact represent a positive outcome for the victim, in some cases it does also provide a platform for the abuser, and incentivises those who are abusive in pursuit of attention to do it again.

For example, one organisation who supports victims of abuse spoke about situations where celebrities or influencers react to the abuse they receive publicly (such as reposting, liking or commenting on it). This abuse will then get a lot of attention – fuelling some abuser's desire for notoriety. The organisation described that abusers may see it as their chance to get acknowledged by a high-profile individual.

Fiona, frequently shares and talks about the abuse they receive online, which invites a lot of welcome support and sympathy from others. While Fiona did not recognise or explicitly talk about this phenomenon themselves, the increased engagement with their profile when they shared this abuse, appeared to lead to more abuse.
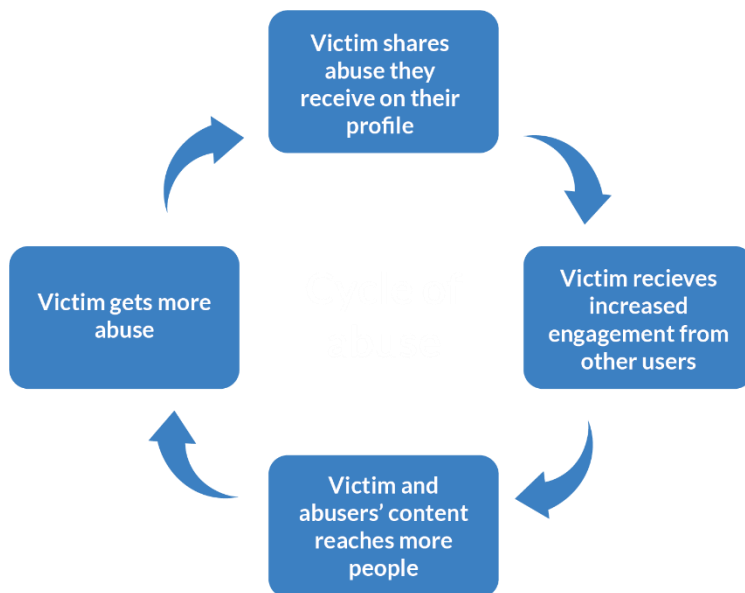
We observed a cycle of abuse happening when victims shared the abuse they received:

Figure 1: Diagram showing cycle of abuse

Other respondents who were victims of abuse were aware of this cycle and dependency dynamic when it comes to sharing abuse, and chose to avoid it:

"I don't do it as much anymore because it can be perceived badly, and also it can cause more problems… I don't even reply to negative comments any more… I try to keep a really low profile on the abuse side of my content… otherwise it perpetuates in different ways… the second-hand abuse that comes from sharing the original abuse… it's exhausting" **Bill, 27**

"I don't like to constantly go on about [negative comments] … I don't want to put those people in the spotlight. I feel like that's a large part of what they want… I don't want to feed into that …" **Jodie, 24**

People move platforms when they feel their views are not tolerated

Some had decided that they did not want to behave within the guidelines of mainstream social media platforms, where they felt their views were not tolerated. Often these people were simply moving to other

areas of the internet where they felt like their views would be more accepted, most commonly Gab, Gettr and BitChute. Others that had been banned from Twitter and other major platforms had also moved to these platforms as a result.

For example, Alan described a gradual, creeping censoriousness that had been building over the years that he was using Twitter. He saw more and more people getting banned or suspended over what he saw as more and more innocuous or inoffensive things. Eventually, he was suspended for something he does not remember posting – but assumes it might have been around the Gender Recognition Reform Bill in Scotland. He had already become disillusioned with Twitter, and this pushed him over the edge. So instead of waiting for his suspension to lift, he left Twitter and migrated over to a platform he saw a lot of people talking about and moving to: Gettr.

"A lot of people involved in debate were being thrown off Twitter…for dubious reasons. I was getting fed up with it anyway. That's why you're seeing people go on Gettr and others. Gettr seems to be the one people are actually going to, because they felt they were being excluded on Twitter." **Alan 60**

Most of the research participants on these platforms who were perpetrators of abuse had, at some point, been banned or suspended from one of the major platforms. Not always knowing what they were banned for, it is hard to know whether they were being 'abusive' or not. What seems to be the case is that, in almost all scenarios, they did not think they were being abusive, and they thought the platform was being overly censorious.

We also cannot say whether some of these people were moving to more 'tolerant' platforms in order to be able to be abusive, or whether they were doing so to feel less censored and more able to express their views.

From observation of the content on these platforms, abusive content is far more prevalent than on mainstream platforms like Facebook and Twitter. This suggests that people migrating to these platforms will inevitably be more exposed to abusive content. This could in turn lead to those people becoming more entrenched in their views or becoming more abusive – however this is not possible to confirm without a longitudinal tracking study.

Section 7

Summary and implications

- There are a range of factors, separate or in addition to anonymity which appear to have an impact on people's abusive online behaviours. These include the disposability of accounts and platform design and reward mechanisms which may incentivise abuse. It is crucial that these factors are considered alongside anonymity when assessing how to reduce abuse online.
- A perception of 'unfair censorship' may lead people towards platforms where they are exposed to content which is more likely to be abusive. This could potentially result in people developing attitudes or exhibiting behaviours which could potentially lead to more harm in the long term. However, to test this hypothesis, a longitudinal tracking study would be required.

Section 8
Conclusions

These conclusions are based on the primary research, which set out to address gaps in previous studies on this subject.

Isolating the role anonymity plays in abusive behaviour is impossible in practice

There appears to be a relationship between anonymity and abuse in some cases but not others. Even where it looks as though there is a link, isolating the role that anonymity plays in facilitating or magnifying abuse is practically impossible – there are always other factors at play as well. It is difficult to disentangle which of these has the greatest impact, and the impact is likely to change depending on the circumstances.

Even with a much larger sample than this qualitative research allowed for, it would be incredibly difficult to determine a causal link between anonymity and abuse given the differing influence of platform design, the behaviour of other users, and the sensibilities of those involved.

Anonymous abuse is not necessarily experienced as being 'worse' than abuse from someone who is identifiable

Some respondents felt that anonymous abuse that was angry and reactionary was easier to brush off or ignore than if it had been from someone identifiable. Some also described that receiving content (such as threatening direct messages) that were anonymous, was more distressing than if it was identifiable, because they could not assess how serious a threat it was.

Platforms are finding and removing significant amounts of abusive online content

In an effort to identify victims and perpetrators of abuse, we searched for abusive content online, though found it difficult to find examples of online abuse, particularly on the mainstream social media platforms. This was not a systematic search or measurement study, as finding online abuse was not the main purpose of the research, but rather an observation.

Many perpetrators and victims of abuse spoke about and showed us screenshots of content that had since been deleted by the platform. This suggests platforms are effectively locating and removing significant amounts of abusive content online. We also know from interviewing perpetrators that platforms are banning or sanctioning users who post abusive content.

Aside from the platform deleting posts and banning users, the onus was often on victims to use tools to tackle abuse. Victims could block accounts that were targeting them, set their own account to private or, on Twitter, select key words which would mean any posts, comments or messages containing them would not be visible. Many of the respondents we spoke to who relied on their public profiles and received huge volumes of messages and comments, felt these tools were ineffective, and would limit their ability to engage with others online in the way they wanted to.

Measures to remove or report abuse also do not stop more abusive content from being posted – and the ease at which people can set up a new account and post again means abuse can be repeated or continue.

Ease of creating new accounts (both anonymous and identifiable) shapes abusive behaviour

The ease with which people can rapidly create new profiles after being banned or suspended from social media platforms appeared to be having a big impact on the volume of abuse on platforms.

Many victims we spoke to thought they were being targeted repeatedly by the same abusers using different accounts set up for the purpose. They felt the ease and speed with which people could create multiple disposable or 'burner' accounts facilitated abuse, as banning or blocking abusive users had little impact on stopping them from posting again.

The fact that people could set up accounts anonymously was one of several factors that made it easy to set up burner accounts. However, it was not the only factor. Removing anonymity would make it more difficult to set up disposable accounts, but there are other ways this could be made more difficult too, which may have fewer negative consequences for other users. For example, placing time limits or restrictions on posting from new accounts, or ensuring users are always identifiable to the platform.

Platforms' enforcement of guidelines or tolerance of abuse influence where people post abuse

Perpetrators of abuse were moving from platforms where abuse was not tolerated (e.g. Twitter) – to platforms where it was (e.g. Gab). It was not the degree of anonymity Gab allowed that seemed to encourage these respondents to move there, but the belief that their views and opinions were tolerated (by the platform and the community) and their accounts would not be suspended.

The benefits of removing anonymity could be achieved in other ways

Few perpetrators of abuse themselves told us they would no longer behave abusively if they were not able to be anonymous online.

We have seen in this research that removing the ability to be anonymous online would likely increase the friction involved in creating accounts online, make it more difficult to set up or post from disposable accounts and make it easier in some cases for platforms to enforce their community guidelines. Yet there are other ways in which these benefits could be achieved, which do not involve restricting anonymity online.

Limiting or banning anonymity would have unintended consequences that would be negative for some people

Making it harder for users to be anonymous could inadvertently expose vulnerable people, or prevent them from having accounts on platforms altogether.

This could risk creating a two-tier internet where certain people, often more likely those at risk of harm or with certain protected characteristics, would be unable to access or make the most of important parts of the online world.

However, many individuals benefiting from anonymity were more concerned about remaining anonymous to other users, rather than platforms themselves.

Anonymity is often an outcome of abusive behaviour, not a precursor

This research has documented a range of online abuse from accounts with varying levels of anonymity. However, most perpetrators of abuse within the sample had not become abusive as a result of the ability to be anonymous.

Instead, they were more likely to have first been abusive while using their real identities, and had only become anonymous once their original profiles had been blocked or banned in an effort to be able to continue to post.

Platforms' design can inadvertently reward abusive behaviours

The reward mechanisms that are designed into social media platforms – likes, follower counts, etc. – incentivise certain behaviours. Many of the perpetrators of abuse we spoke to were seeking attention or validation for their opinions. Some felt incentivised to post shocking or inflammatory content to get attention, and were rewarded with 'likes' or more followers, and comments or people sharing their content – whether or not they agreed with it.

There were perverse incentives for victims to enable the cycles of abuse

Some of the victims often shared the abuse they received, to 'call out' the abusers or fight back. This led in many cases to an outpouring of support and attention. The victim could see this as a positive outcome, but it does also provide a larger audience for the abuser, and could incentivise those who are abusive and in pursuit of attention to repeat, copy or escalate their behaviour.

For some victims it was clear this behaviour of re-posting abusive content was resulting in more abuse. Two of the people we spoke to who had large followings felt some people were reliant on the abuse they received, as it was part of their online identity, and meant they received outpourings of support from many of their followers.

Recommendations for future research and work

This research has highlighted several opportunities for future research and work in this area.

Future research should be done to:

- Test measures that would make it more difficult to create and / or post from disposable accounts, and explore the impact on abusive behaviour

- Explore the implications of users having to be identifiable to platforms, but not to other users, on abusive behaviour. This would mean users could be better held to account for the content they post without compromising the benefits to some users of being anonymous to other users.

- Investigate the effects over time when perpetrators of abuse move to platforms which they feel are more tolerant of abusive behaviour, for example exploring whether posting and consuming content in those online spaces entrench particular behaviours or attitudes.

This research also has implications for work being done to reduce online abuse:

- More work should be done to ensure people who are abusive or have been suspended from a platform are not able to create new, anonymous, disposable accounts.

- Measures taken to reduce online abuse should protect people's ability to create profiles in which they are anonymous to other users.

Annex: Full rapid evidence review

Introduction

There is much discussion around the role anonymity plays in the prevalence and severity of online abuse. The issue of online anonymity also plays into a larger ongoing debate relating to the balance of freedom and protection online. Each perspective comes with a range of evidence, significant lobbying power, valid concerns, and conflicting policy demands.

DCMS commissioned work to explore these perspectives and assist in finding a balance between the positives and negatives of anonymity within online networks. This aim of this research was to support DCMS in its efforts to reduce abuse and its harmful impacts, without overly reducing the positives of unrestricted, anonymous online participation.

This document, outlining learnings from a rapid evidence review which explores evidence from both sides of the debate, is part of that work. The findings from this document will be combined with insights from expert interviews and qualitative interviews with victims and perpetrators of online abuse, as well as people benefiting from online anonymity.

This rapid evidence review set out to explore:

- Definitions of abuse and anonymity
- The benefits of anonymity for certain user groups
- Existing verification methods and measures to deal with online abuse by platforms
- Existing research into the role anonymity plays in facilitating online abuse, and causal links
- International approaches to dealing with anonymity and abuse

The studies reviewed for this report were largely published in scientific journals, meaning they have undergone some form of peer review.

NB. This was not a systematic literature review and did not cover all existing evidence relating to the link between abuse and anonymity

Examples of search terms used throughout the process:

The studies reviewed for this report were largely published in scientific journals, meaning they have undergone some form of peer review.

A comprehensive list of search terms were not recorded as this was a rapid review, however, search terms broadly covered:

Online abuse				
Abuse	Online abuse definition	Online abuse prevalence	Recipients of online abuse	Cyberhate
Trolling	Online hate	Online harms	Impact of online abuse	Victims of online abuse
Online anonymity				
Anonymity	Online anonymity definition	Platforms and anonymity	Social media and anonymity	Benefits of anonymity
A search for social media platforms' approaches to anonymity				
The relationship between anonymity and abuse				
Anonymous versus known abuse	Link between anonymity and abuse	What factors lead to online abuse	The Disinhibition Effect	The impact of anonymity online

Anonymity and social media	Cyberbullying and anonymity	Online abuse and anonymity	Online hate and anonymity	
Existing policy recommendations				
Online Safety Bill call for evidence	Anonymity and abuse online	Anonymity and abuse online policy recommendations	*A search of charities and lobbying organisations in this space*	
International approaches to dealing with anonymity and abuse				
Online anonymity legislation	International approaches to anonymity	International approaches to abuse	International online safety	*A search for countries known to have put in place measures in this space*

Searches took place on Google Scholar, EBSCO Discovery Service (for academic literature) and Google Search (for journalistic articles, research commissioned by charities, and policy recommendations).

In addition to using search terms to find literature, much of the literature explored was found through citations and links from other papers, i.e. exploring literature cited across multiple evidence reviews.

Online abuse

Defining online abuse

Existing research usually separates online abuse into two main types[65]:

- Abuse directed against a group, usually called 'hate speech'
  - "[Hate speech] broadly includes negative textual, visual or audio-based rhetoric that attacks, abuses, insults, harasses, intimidates, and incites discrimination or violence against an individual or group due to their race, ethnicity, gender, religion, sexual orientation or disability[66]".
- Abuse directed against an individual, usually called 'harassment' or 'cyberbullying'
  - A definition for online harassment is: "aggressive, intentional acts carried out by a group or individual, using electronic forms of contact against an individual"[67]

The definitions above provide a useful summary, but there are no universally accepted or consistently used definitions for online hate speech or online harassment—as highlighted in Davidson et al.'s Rapid Evidence Assessment of Adult Online Hate, Harassment and Abuse[68] and the Turing Institute's review into the prevalence of Online Abuse.[69]

Definitions tend to focus on describing different types of abusive behaviour and the intent behind this behaviour (e.g. to intimidate or insult).

The main types of online harassment encountered according to Davidson et al., 2019[70] are:

- Offensive name calling
- Purposeful embarrassment
- Physical threats
- Sustained harassment
- Stalking
- Sexual harassment

Other abusive behaviours include:

- Trolling – "deliberately posting offensive, upsetting or inflammatory comments online in an attempt to hurt and provoke a response"[71]
- Pile-ons – many individuals, acting separately, send messages that are harassing in nature to a victim. When taken together these messages may cause alarm or distress, even though each individual message may not reach a criminal threshold[72]

---

[65] Vidgen, B., Margetts, H. and Harris, A., 2019. *How much online abuse is there? A systematic review of evidence for the UK Policy Briefing – Full Report.* Public Policy Programme Hate Speech: Measures and Counter Measures. [online] The Alan Turing Institute.

[66] Davidson, J., Livingstone, S., Jenkins, S., Gekoski, A., Choak, C., Ike, T. and Phillips, K., 2019. *Adult Online Hate, Harassment and Abuse; A Rapid Evidence Assessment.* [online]

[67] Dadvar et al., 2013 as cited in Vidgen, B., Margetts, H. and Harris, A., 2019.

[68] Davidson, J., Livingstone, S., Jenkins, S., Gekoski, A., Choak, C., Ike, T. and Phillips, K., 2019. *Adult Online Hate, Harassment and Abuse; A Rapid Evidence Assessment.* [online]

[69] Vidgen, B., Margetts, H. and Harris, A., 2019. *How much online abuse is there? A systematic review of evidence for the UK Policy Briefing – Full Report.* Public Policy Programme Hate Speech: Measures and Counter Measures. [online] The Alan Turing Institute.

[70] Davidson, J., Livingstone, S., Jenkins, S., Gekoski, A., Choak, C., Ike, T. and Phillips, K., 2019. *Adult Online Hate, Harassment and Abuse; A Rapid Evidence Assessment.* [online]

[71] Stop Online Abuse, 2022. *What is online abuse?* [online]

[72] Law Commission, 2020. *Harmful Online Communications: The Criminal Offences - A Consultation Paper.* [online]

- Doxxing – publishing personal information about someone else online
- False profiles – fake social media accounts set up with the intention of damaging a person's reputation[73]

Challenges in defining online abuse

Online abuse is a broad category containing many different types of behaviour, which are likely to have different outcomes depending on the severity, type and recipient.

DCMS' Online Harms Feasibility Study[74] used the framework of hazards, risks and harms to explain the different outcomes a piece of online content, interaction or design feature (all hazards) can have. Whether hazards become harmful, or in this context are seen/felt to be abusive to an individual or not, is highly dependent on risk factors, including the context in which the hazard occurred, the individual themselves and the level of exposure to a hazard.

Definitions of online abuse are also highly subjective and dependent on interpretations of intent and impact. Imprecise definitions of abuse can lead to discussions of online behaviour conflating *incivility* or *impoliteness* with *abusiveness*. Incivility has been applied to a range of behaviours, including "aggressive commenting in threads, incensed discussion and rude critiques […] outrageous claims, hate speech and harassment"[75]. Therefore, whilst an uncivil or impolite online environment can cause harm, conflating incivility with abusiveness ignores the democratic merit of robust and heated discussion, and the varying norms of communication across groups.[76] [77]

How widespread is online abuse?

The prevalence of abuse online has been predominantly measured using **surveys** and **measurement studies** (in which academics analyse samples of social media data, e.g. a dataset of Tweets, to detect the prevalence of potentially harmful content[74]). **Platform transparency reports** have also been used to indicate prevalence of online abuse.

The limitations of these methods for measuring the prevalence of abuse have been explored in detail within DCMS' Online Harms Feasibility Study.[74]

The study found that while a variety of measures and sources exist, they do not provide an accurate or consistent enough picture of online harm to reliably measure prevalence or impact over time—there is no 'ready-made' solution—and there are numerous challenges to measuring online harm well. Collecting the breadth and depth of data required to make accurate estimates of the volume and impact of online harm is a significant undertaking. Broadly, the study finds that the challenges / barriers to measuring prevalence fall into several categories:

#1. Most potential measures / sources were not designed with the intention of being used as a measure of online harm

#2. Technical limitations of research methods employed (as explored later in this evidence review)

#3. Lack of consistent definitions and granularity

---

[73]Cybersmile, 2022. *What Is Cyberbullying? – Cybersmile*. [online] Cybersmile.org.

[74] DCMS, 2021. *Online Harms Feasibility Study*. Unpublished (to be published in 2022)

[75] Antoci, A., Delfino, A., Paglieri, F., Panebianco, F. and Sabatini, F., 2016. Civility vs. Incivility in Online Social Interactions: An Evolutionary Approach. *PLOS ONE*, 11(11).

[76] Papacharissi, Z. 2004. Democracy online: civility, politeness and the democratic potential of online discussion groups. *New Media & Society* 6(2):259-283;

[77] Muddiman, A. 2017. Personal and Public Levels of Political Incivility. *International Journal of Communication* 11

The Alan Turing Institute (2019) conducted a systematic review of evidence regarding the prevalence of online abuse in the UK[78]. At the time of developing the Online Harms Feasibility Study in 2021[79], this study was the most comprehensive review of available literature. This research reviewed surveys, measurement studies and platform transparency reports to estimate the prevalence of online abuse. A summary of findings is shown in the table below.

Study	Method	Finding
[How much online abuse is there?](#) **Alan Turing Institute** 2019 *See report for list of studies included in this review and their individual findings*	**Review of surveys**	The report estimates 30-40% of people in the UK are exposed to abuse online, with 10-20% targeted by abusive content online. *The first figure is based on results from the Oxford Internet Study. This asks "Have you seen cruel or hateful comments or images posted online?" It does not ask questions around the regularity of seeing online abuse – the respondents could have seen abuse once or countless times. It also does not show how many people have personally received abuse.* *The study also asks "have you received obscene or abusive emails?", which is fairly limited in scope.*
	Review of measurement studies	Measurement studies show that the prevalence of abuse is less than 1% on mainstream platforms. This is predominantly based on studies of abuse on Twitter, due to the availability of data. Less mainstream sites such as 4chan, 8chan, and Stormfront have been shown to have a higher prevalence of content considered to be hateful – approximately 5-8%. However, most research in this domain has focused on hate speech analysis and there is a lack of research into person-directed abuse, such as harassment. *The Alan Turing Institute recognises that measurement studies are often unrepresentative or not UK-focused, and that AI tools often have a high level of error when applied in 'the wild'.*
	Review of platform transparency reports	The report estimates that the amount of abuse on mainstream platforms which is serious enough for the platforms to action is ~0.001%, although this is speculative as platforms do not share how much total content they host. *The Alan Turing Institute recognises the limitations of transparency reports – including that they are often not UK-focused, and that platforms can only report on the abuse they have found, not that exists in total.*

More recent evidence from a survey of US adults in 2020 highlighted that 41% of adults had personally experienced some form of 'online harassment'. This has increased from 25% in 2014[80].

---

[78] Vidgen, B., Margetts, H. and Harris, A., 2019. *[How much online abuse is there? A systematic review of evidence for the UK Policy Briefing – Full Report.](#)* Public Policy Programme Hate Speech: Measures and Counter Measures. [online] The Alan Turing Institute.

[79] DCMS, 2021. *Online Harms Feasibility Study*. Unpublished (to be published in 2022)

[80] Pew Research Center: Internet, Science & Tech. 2021. *[The State of Online Harassment.](#)* [online]

Which users are more likely to experience online abuse?

There is a convincing amount of literature that explores whether certain groups are more likely to receive online abuse than others based on:

- **Ethnicity.** From a review of UK studies on online abuse, those from Black and 'Other' ethnicities were more likely to report being targeted with online abuse compared to White or Asian people[81]
- **Sexuality.** Lesbian, gay or bisexual adults in the US were more likely to report receiving online harassment compared to straight adults (51% of Lesbian, Gay and Bisexual (LGB) adults were targeted with more severe forms of online abuse compared to 23% of straight adults)[82]
- **Disabilities.** From a review of UK studies on online abuse, those with disabilities reported seeing more online abuse than those without disabilities[83]
- **Gender.** A review of studies in Scotland and the UK exploring the prevalence of online abuse for trans individuals found that between 34%-52% of trans individuals had experienced online abuse[84]
- **High profile individuals.** Qualitative research with UK MPs[85], as well as measurement studies[86], have highlighted the online abuse received by MPs. Other high-profile individuals, such as footballers and influencers, have also been highlighted as key targets for online abuse[87] [88]

However, few of the existing studies of UK adults explored in the examples above cover a nationally representative population and instead focus on specific groups (e.g. experiences of MPs). This means it is not possible to break down data relating to experiences of online abuse across different demographic groups, making it hard to compare prevalence rates among UK adults.

Impact of online abuse

Some literature has highlighted the impact online abuse can have on:[89] [90] [91]

- **Emotional and mental health.** For example, people who have experienced online abuse report suffering from anxiety, depression, and panic attacks. Other reports include lower self-esteem, and feelings of anger and shame
- **Fears for physical safety.** People report fearing for their physical safety offline
- **Changes to online activity**. For example, some people report restricting the use of their social media account, or deleting their accounts all together
- **Impacts on friends and family**. Friends and families of victims also report being affected – often becoming concerned about the victim's safety, as well as being targeted or threatened themselves

However, it is worth noting that many studies exploring the impact of online abuse do not specify the *type* of online abuse someone has received and explore how this relates to impact. As such, abusiveness is imagined as

---

[81] Vidgen, B., Margetts, H. and Harris, A., 2019. *How much online abuse is there? A systematic review of evidence for the UK Policy Briefing – Full Report.* Public Policy Programme Hate Speech: Measures and Counter Measures. [online] The Alan Turing Institute.

[82] Pew Research Center: Internet, Science & Tech. 2021. *The State of Online Harassment.* [online]

[83] Vidgen, B., Margetts, H. and Harris, A., 2019. *How much online abuse is there? A systematic review of evidence for the UK Policy Briefing – Full Report.* Public Policy Programme Hate Speech: Measures and Counter Measures. [online] The Alan Turing Institute.

[84] Ibid.

[85] Committee on Standards in Public Life, 2021. *Intimidation in Public Life: A Review by the Committee on Standards in Public Life*.

[86] Farrell, T., Bakir, M. and Bontcheva, K., 2021. *MP Twitter Engagement and Abuse Post-first COVID-19 Lockdown in the UK: White Paper.* [online]

[87] UK Parliament Committees. 2020. *Anton Ferdinand, Lianne Sanderson and Marvin Sordell give evidence on online racist abuse.* [online]

[88] BBC News. 2021. *Social media influencers face relentless abuse, MPs are told.* [online]

[89] Hubbard, L., 2020. *Online Hate Crime Report: Challenging online homophobia, biphobia and transphobia*. Galop.

[90] End Violence Against Women and Glitch, 2020. *The Ripple Effect: Covid-19 and the Epidemic of Online Abuse*.

[91] Amnesty International. 2017. *Amnesty reveals alarming impact of online abuse against women*. [online]

one behaviour, rather than a set of discrete behaviours. For example, it would be valuable to distinguish between the impact of; sustained trolling from a specific user and a one-off pile-on; sexual harassment and physical threats; or abuse from anonymous accounts and known accounts. This would better inform actions and policies that can target specific abusive behaviours.

Furthermore, the impact of abuse can vary based on the recipient. What can be considered 'threatening', 'embarrassing' or 'attacking' will likely vary from person to person and is highly dependent on the personal characteristics of the recipient, as well as the context in which it occurs. Abuse could also have an immediate and obvious impact on some individuals, while others may not recognise any impact, particularly in the short term. This further complicates the ability to measure the impact of abuse, as many studies rely on self-reporting by recipients.[92]

Is anonymous abuse experienced differently to 'known' abuse?

There appears to be limited research on the impact of known versus anonymous abuse on victims. In a qualitative study with 25 young people aged 15 to 24 on risk factors associated with the impact and severity of online bullying, participants reflected on what impact the anonymity of the perpetrator had on feelings of loneliness, fear and worry.[93] Some respondents felt an increased impact if their bully was anonymous as there was a fear in the unknown: *"I felt victimized and scared because I did not know who'd done it"*. However, others felt more impact when their bully was known to them: *"That had a huge effect on me because that person used to mean so much to me."* This suggests that while the type of impact may be different, it is difficult to determine whether anonymous abuse has more impact on the recipient: this will vary greatly depending on the individual, context and content of the abuse.

Current methods platforms use to combat online abuse

Social media platforms currently use a range of tools which aim to combat online abuse.[94] [95] [96] [97] These fall into three broad categories:

- setting rules and guidelines about online abuse
- identifying online abuse
- acting on online abuse

This table separates the tools used by platforms into these three categories, and adds descriptions as to how they operate:

Category	Tool	Detail
Setting guidelines	Platform guidelines	Community guidelines or standards outline what is and is not allowed on a platform
Identifying online abuse	User reporting tools	Functions such as a reporting button, which allow users to flag specific content to the platform. This is often then reviewed by AI or human moderators to decide if it needs to be removed from the platform

[92] DCMS, 2021. *Online Harms Feasibility Study*. Unpublished (to be published in 2022)

[93] Dredge, R., Gleeson, J. and de la Piedad Garcia, X., 2014. Risk Factors Associated with Impact Severity of Cyberbullying Victimization: A Qualitative Study of Adolescent Online Social Networking. *Cyberpsychology, Behavior, and Social Networking*, 17(5), pp.287-291.

[94] Facebook Meta, 2022. *Transparency Center.* [online]

[95] Twitter. 2022. *Help Center.* [online]

[96] TikTok. 2022. *Safety & privacy controls | TikTok.* [online]

[97] Reddit. 2022. *Content Policy - Reddit.* [online]

	User (volunteer) moderators	A user volunteers to moderate content within a small section of the platform. For example, a Reddit user can volunteer their time to moderate one of Reddit's communities. They are then responsible for setting and monitoring the rules of that community, which may go above and beyond Reddit's platform-wide guidelines
	Platform (professional) moderators	Official moderators employed by the platform to review content flagged by users or its automated tools
	Automated tools	AI tools which automatically pick up and block certain types of content, such as particular images, words or phrases
	Policy officials	Provide support in wider trust and safety teams. They help train moderators and improve policy when new violations appear
Acting on online abuse	Removing content	Removing individual posts Removing pages / groups / threads This can be enacted by both professional and volunteer moderators
	Limiting the prominence of content	For example: preventing posts from appearing in 'explore' feeds or in recommendations; limiting the extent to which content can be shared
	Adding warnings or flags to sensitive content	Providing warning screens or flags to content that some users may find sensitive
	Providing 'warnings' to accounts	Issuing a 'warning' or 'strike' to accounts sharing content that goes against community guidelines
	Muting or restricting accounts	Restricting accounts from creating content for a certain period of time. For example, a Facebook user with three 'strikes' will receive a three day restriction from creating content[98]
	Permanently blocking or removing accounts	A user can decide to permanently block an account. This may also include the ability to block new accounts set up with the same details A platform can decide to permanently remove an account from their platform
	Working with law enforcement to support investigations	Some platforms have dedicated law enforcement teams to support this Platforms can provide law enforcement with identifiable information for account holders, if they are in possession of this information

---

[98]Facebook Meta. 2022. *Restricting Accounts | Transparency Center.* [online]

Online anonymity

Online anonymity cannot be understood in binary terms. For example, a platform could allow a user to be entirely anonymous – unidentifiable to both the platform and to other users – or it might require users to be identifiable to the platform but allowed to use a pseudonym – concealing their identity from other users.

Anonymity is also utilised for a range of reasons. By freeing people from the social norms and reputational risk that guide their behaviour *offline*, anonymity might facilitate people posting abusive content, but it might also be used by people who want to be themselves online in a way they feel they cannot be offline. The versatility of anonymity is precisely what makes policy decisions so complex.

Platforms and anonymity

As noted in the table below (accurate at the time of writing), most popular social media sites enable a level of pseudonymity, except for Facebook and LinkedIn which encourage 'real name' policies. However, these policies are difficult to enforce, and have received criticism for cultural bias. Moreover, they are often not able to determine that the *real-sounding* name you use is actually *your* name.

Upon signup to all the sites in this table, a user is required to provide an email address which is verified by using click-through verification via email. This is often known as an authentication process. There is no limit nor regulation on the number of 'burner' or fake email addresses a user can make and use to sign up to any of these platforms, nor does this method of email verification mean you are necessarily identifiable by the platform. Some platforms collect IP addresses at sign-up, but the use of VPN and plugins which redirect or bounce your IP address offer an easy way to get around this.[99]

| Platform | Levels of anonymity: | | |
	Real Name Policy	Age requirement	Contact details
LinkedIn	Yes *If an account is reported as pseudonymous or otherwise restricted, the user is locked out and the platform requests verification in the form of official ID*	16 (previously 13) *But, birthdate is not requested at sign up[100]*	Valid email address
Facebook	Yes *But the policy has been 'watered down' after facing backlash[101]* *If an account is reported as pseudonymous, the user is locked out and the platform requests verification in the form of official ID*	13+ (higher in some countries) *No verification of birthdate at sign up. But, other users can report underage accounts and AI technology is being used to detect underage accounts[102]*	Valid email address *or* mobile number

[99]Martin, J., 2020. *Does a VPN make you anonymous?.* [online] Tech Advisor.

[100] LinkedIn. 2022. *User Agreement*. [online]

[101]York, J., 2014. *Facebook's 'real names' policy is legal, but it's also problematic for free speech | Jillian C York.* [online] The Guardian.

[102] Facebook. 2021. *How Do We Know Someone Is Old Enough to Use Our Apps?* [online]

Instagram	No	13+ (higher in some countries) *No verification of birthdate at sign up. But, other users can report underage accounts and AI technology is being used to detect underage accounts* *ID requests may be made to verify age when appeals are made by users who have been suspended due to being underage[103]*	Valid email address *or* mobile number
Twitter	No	13 *If an account is reported as underage, "the reported account owner's age must be reasonably verifiable as underage in order to take action"[104]*	Valid email address *or* mobile number
TikTok	No	13 *Birthdate requested at signup, no age verification mechanism* *But, other users can report underage accounts and AI technology is being used to detect underage accounts[105]*	Valid email address *or* mobile number
Tumblr	No	13 *Birthdate is not requested at signup, no age verification mechanism[106]*	Valid email address
Reddit	No *Real names are discouraged*	13 *Birthdate is not requested at signup, no age verification mechanism[107]*	Valid email address
Snapchat	No	13 *Birthdate requested at signup, no age verification mechanism[108]*	Valid email address *or* mobile number
YouTube	No	13 *Birthdate requested at sign up* *New users must provide age verification (ID or credit card)*	Valid email address

---

[103] Instagram. 2022. *Why is Instagram asking for my birthday?* [online]

[104] Twitter. 2022. *Staying safe on Twitter and sensitive content.* [online]

[105] TikTok. 2022. *Guardian's Guide.* [online]

[106] Tumblr. 2022. *Your Age on Tumblr.* [online]

[107] Reddit. 2022. *Content Policy - Reddit.* [online]

[108] Cuthbertson, A., 2019. *Snapchat admits its age verification system does not work.* [online] Independent.

		if/when accessing age-restricted content[109]	

Who benefits from anonymity?

Discussions regarding the benefits of anonymity online for particular people or groups of people is largely anecdotal or based on secondary research into related topics. It focuses on those who, due to their offline identities, occupations, social networks, or communities, may be at risk both online and offline if their identities were not concealed. Much of this conversation in the UK was inspired by the emergence of the idea of introducing real-name policies and ID verification across social media platforms, following the racist abuse of prominent footballers during the Euro 2020 tournament.[110] It was purported that reduced anonymity online by the introduction of real-name policies would exclude those unable to provide ID verification and unfairly expose others.[111] [112]

LGBTQI+[113] people, victims of domestic abuse[114] [115], those with taboo or alternative opinions, beliefs or lifestyles[116], sex workers[117], undocumented migrants[118], teachers, public figures, whistle-blowers and activists[119] may all be considered to benefit from anonymity online. It could be said that the ability to remain anonymous to other users allows people to build connections, share stories, engage in conversation, find companionship and support and more generally just 'exist' online with a greater degree of safety, confidence and comfort.

However, the type of anonymity to which these conversations refer is ambiguous and ill-defined. Although there is a distinct conversation related to the protection anonymity provides from other online users, the shape or 'level' of that anonymity is unclear. Much of the anecdotal evidence cited above, positioned against the idea of real-name policies, suggests that users and advocates of anonymity care about protecting the freedom to use pseudonyms or anonymous user-facing profiles to create 'safe spaces'[120] and promote freedom of expression. Arguments against the introduction of ID verification tend to focus on whether this would act as a barrier for specific groups to access platforms, such as undocumented minority communities, as well as the possibility of state surveillance which might limit the freedoms of activists, whistle-blowers or those who organise against the government.[121]

More broadly, a range of academic papers explore the wider costs and benefits of anonymity online and point theoretically or evidentially, based on small samples, to some groups of beneficiaries. Some of these debates are evidenced below, collectively addressing anonymity as it relates to a range of human rights and social norms.

Study	Key Findings/Argument	Limitations of study

[109] YouTube Help. 2022. *Age-restricted content.* [online]

[110] Smith, A., 2021. *Calls to end social media anonymity give platforms more power without actually fixing the problem, experts say.* [online] Independent.

[111] Bold, B., 2021. *Should social media users be required to prove their ID to stamp out online abuse?.* [online] Campaign.

[112] Dyer, H., 2021. *Online abuse: banning anonymous social media accounts is not the answer.* [online] The Conversation.

[113] St James, E., 2020. *Trans Twitter and the beauty of online anonymity.* [online] Vox.

[114] Baker, S., 2016. *Why Online Anonymity is Critical for Women - Women's Media Center.* [online] Womens Media Centre.

[115] SM Kee, J., 2007. *CULTIVATING VIOLENCE THROUGH TECHNOLOGY? Exploring the Connections between Information Communication Technologies (ICT) and Violence Against Women (VAW).* [online] APC WNSP.

[116] Noman, H., 2015. Arab Religious Skeptics Online: Anonymity, Autonomy, and Discourse in a Hostile Environment. *SSRN Electronic Journal.*

[117] Andrews, I., Undated. *A Brief Guide to Online Privace and Security for Sex Workers.* [online]

[118] Stop Online Abuse, 2022. *What is online abuse?* [online]

[119] Van der Merwe, B., 2021. *Are anonymous accounts responsible for most online abuse?.* [online] New Statesman.

[120] Herrman, J., 2021. *You Anon.* [online]. The New York Times.

[121] Khan, M., 2021. *Using photo ID to verify social media accounts could harm minorities more than it protects them.* [online] iNews.

The Value of Anonymity on the Internet Lim, Zo & Lee. 2011	• This study ran a simulation test based on an opinion diffusion model to test the effects of anonymity on society as a whole • Broadly, the simulation found that anonymity is more useful in closed societies than open ones – in open societies, where anonymity is not a necessity for free discourse, anonymity may reduce credibility • In a closed society, anonymity could facilitate the flow of information between members of different offline groups • The impact of anonymity is extremely context-specific and affected by things like the openness of the society, the level of false information in the society, and balances of social and governmental power	• This simulation uses a very small number of test agents – the results are not generalisable and do not reflect society • The simulation was repeated just 30 times, more simulations with more agents would provide stronger evidence
The positive and negative implications of anonymity in Internet social interactions: "On the Internet, Nobody Knows You're a Dog." Christopherson. 2007	• Outlines a range of contemporary theories about the effects of anonymity and suggest that they are relevant for computer-mediated communications (CMC) • Anonymity can have positive and negative outcomes mediated by how one chooses to use it • The strategic use of anonymity can help empower marginalised communities • Conversely, anonymity can also encourage anti-normative and discriminative behaviour or predatory behaviour • Anonymity can protect people from discrimination based on gender, race, age, ethnicity, physical disability, and attractiveness if their appearance is concealed • The impact of anonymity is also differentiated by gender – it is more common for women to use anonymity as a safety strategy • Concludes that CMC are not necessarily mediated strictly by anonymity but also social norms	• While this paper provides useful theoretical overview and frameworks for assessing the impact of anonymity online, it is not based on original research
Why do people seek anonymity on the Internet? Informing policy and design Kang, Brown and Kiesler. 2013	• Qualitative report based on 44 interviews • Almost half of the participants had used anonymity to attack others online, but the line between the illegality and undesirability of that activity was unclear • Many of the respondents used anonymity to participate in interest or hobby groups where the "norm" was anonymity. The use of anonymity was regulated by the social norm in the online space • Some respondents used anonymity to engage in discussions or hobbies their offline communities may not tolerate • The idea of 'threat' was a motivation in the use of anonymity – i.e. perceived repercussions from their identity being exposed. 'Threats' fell into 5 categories; online predators, organizations, known	• As the qualitative report was based on only 44 interviews, it is not representative of the population of internet users • Study design did now allow for a distinction between cultural and political factors in motivations for anonymity

	others, other users on the site or in the community, and unknown others ● Makes a definition between user-defined anonymity and full anonymity; some respondents differentiated between being anonymous from other users as opposed to the site administrators ● Decisions to be anonymous were informed by previous negative experiences of non-anonymity and boundary management	
Freedom of Expression and the Internet Council of Europe. 2013	● Argues that anonymity and pseudonyms have a long tradition in the press, among authors, critics and for whistleblowing ● Sees anonymity on the internet as fundamental to freedom of expression. Uses a range of other rights to support this claim, such as the right to whistleblowing, freedom of assembly and association, right to education and access to knowledge	● This is a largely policy-focused review of the internet in the context of traditional legal and human rights ideas
The right to freedom of expression and the use of encryption and anonymity in digital communications Association for Progressive Communications. 2015	● Links anonymity online to other human rights and sees online anonymity as fundamental to freedom of expression, especially in repressive environments ● Argues that the right to privacy means that it is reasonable that online users should have the ability to control their privacy online, of which the freedom to use anonymity is a factor ● Anonymity also encourages freedom of association and assembly ● Points to a range of different groups for whom online anonymity aids freedom of association and assembly, particularly online association based on identities or beliefs that are illegal or taboo, like LGBT groups, political opposition, or religious minorities ● Argues that anonymity also provides an enabling environment for people to form relationships and seek support for problems that have a social stigma like drug addiction, illnesses such as HIV/AIDS, or sexual abuse	● This paper is founded on legal frameworks, not original research
Arab Religious Skeptics Online: Anonymity, Autonomy, and Discourse in a Hostile Environment Noman. 2015	● Highlights the importance of anonymity for vulnerable people or those with beliefs or opinions which may put them at risk within their offline communities ● Argues that anonymity online can empower marginalized communities ● Explores the freedom anonymity online affords to Arab religious sceptics in environments where their views may carry extreme offline repercussions	● Although the study is detailed, it is based on a review of conversation on only three Arab online forums. Analysis of more forums would provide may have provided more reliable findings
Anonymity, Democracy, and Cyberspace Akdeniz. 2002	● Anonymity is 'socially useful' ● Is a vital tool for the preservation of political speech, freedom of expression and privacy ● Closely related to free speech as a concept	● Is – at the time of writing – 20 years' old. The studies explored will likely

	• Anonymity is important to free speech, privacy and people's participation in online support groups/services (Samaritans, Childline, AA, NA) • Despite these benefits, anonymity also encourages cybercrime	be out of date as online spaces have evolved significantly in this time • It is not based on original research
In defence of anonymity on the internet Marsh. 2007	• Anonymity is linked to freedom of speech and privacy as constitutional rights • Anonymity encourages protection of political speech in democracies • Anonymity may encourage hate speech as it erases accountability but remains important for privacy and constitutional freedoms. Sees it as a contentious 'payoff'	• A US study, which needs to be taken in the context of attitudes towards freedom of speech in that country • Is a secondary research paper – relying on external sources to make the argument for the protection of anonymity (likely an imbalanced representation of available literature). • Is – at the time of writing – 15 years' old, the studies explored will likely be out of date as online spaces have evolved significantly in this time

The relationship between anonymity and abuse

Online anonymity has been under scrutiny due to its proposed link with abusive behaviours. Broadly, there is a tension between studies that suggest the cloak of anonymity enables users to betray group norms and others that suggest anonymous users maintain the same, if not more, attachment to norms. Other studies suggest alternative factors determine abusiveness. We explore in detail this range of studies in a table within this section.

Many online harms reports are inconclusive on the role of anonymity in abuse, often because they rely on imprecise definitions of *abuse* and *anonymity*. For example, a 2020 YouGov poll found that 83% of British people think the ability to post anonymously makes people "ruder" online. This has been cited by *Clean Up the*

Internet in its arguments that link anonymity with abuse but, as highlighted earlier, it is problematic to conflate abusive behaviour with impoliteness.[122]

Similarly, lots of surveys do not distinguish between *strangers*, *anonymous accounts* and *pseudonymous accounts* when they ask respondents about their experience of online abuse, including those from *The Pew Research Centre*[123], *Amnesty International*[124], *Compassion in Politics*[125] and *Glitch*[126]. Due to the lack of a consistent definition, these surveys cannot be used to definitively demonstrate the relationship between anonymity and abuse.

An assessment of methods

Studies have tried to show causal links between anonymity and abuse with a range of methods, including lab studies, data scraping, surveys, evidence reviews, qualitative studies and transparency reports. The table below outlines some of the strengths and weaknesses of each method.

Method	Strengths of method	Weaknesses of method
Lab studies	● Enables researchers to separate and test influencing factors	● Often only conducted in one online space with distinct norms and might not be replicable ● Difficult to replicate organic communications with controlled experiments
Data scraping/social listening/discourse analysis/content analysis	● Enables analysis at scale ● Allows for trends/patterns to be detected and monitored over time	● Coding schema can be ambiguous, for example when definitions of 'incivility' and 'abuse' are too broad ● Does not account for differing social norms between online spaces
Social Network Analysis	● Allows for the detection of patterns of influence of users (nodes) in networks ● Allows for trends/patterns to be detected and monitored over time	● Focusing on user influence may miss important contextual factors of online spaces which regulate communication ● Can have a time-consuming and intensive quantitative analysis process ● May not be able to account appropriately for multiple digital accounts made by a single person
Regression Analysis	● Can be used to map the links between variables (usually from survey data)	● Dependent on the quality of the data set
Surveys (of quantitative or qualitative data)	● Can be used to capture experiences of abuse where anonymity was or was not a factor from a large number of people	● Self-report relies on subjective understanding of 'what counts' and cannot be used to gauge harms that people do not recognise have happened to them ● Sampling presents a significant challenge. 'Hard to reach' groups are often underrepresented and challenging to engage via standard sampling approaches,

---

[122]Babbs, D., 2020. *New opinion poll: 83% of Brits think anonymity makes people ruder online.* [online] Clean up the Internet.

[123] Pew Research Center: Internet, Science & Tech. 2021. *The State of Online Harassment.* [online]

[124] Amnesty International. 2017. *Amnesty reveals alarming impact of online abuse against women.* [online]

[125] Compassion in Politics. 2021. *Three-quarters of those experiencing online abuse say it comes from anonymous accounts.* [online]

[126] End Violence Against Women and Glitch, 2020. *The Ripple Effect: Covid-19 and the Epidemic of Online Abuse*.

	• They can be relatively easy and inexpensive to run and replicate over time • As long as sample is available, they can be used to collect data from specific populations (e.g. a representative survey of adult social media users)	such as using commercial research panels. Reaching truly representative samples is generally extremely difficult and acknowledging the limits and in-built biases of a sample is fundamental to interpreting data accurately • There are only so many questions that a survey can include, so capturing an appropriate level of depth and context around any one online experience will always be challenging
Evidence reviews	• Can provide a thorough understanding of the current research landscape	• Rely on existing data and have no influence over quality of studies and the topics explored
Qualitative studies	• The in-depth nature of qualitative research provides a level of data/insight that is not possible to gain from other methods • Additional techniques (e.g. validated self-report, behavioural tracking) can provide more objective insights into behaviour • The targeted nature and small sample sizes make it far easier to explore hard-to reach groups and sensitive topics	• Small sample sizes mean that findings are not necessarily representative of the wider population • It can be particularly difficult to understand the scale of the problem • As with other methods, the results can be dominated by the experiences of a few with the loudest voices
Transparency reports	• Platforms have access to accurate data on whether abusive content that has been identified was posted by an anonymous account or not • Law enforcement requests/responses within transparency reporting helps to uncover anonymous abusers	• Only provides information on content which has been identified by the platform • If identified by AI it will not have taken into account the context of the post which may mean it is more or less abusive • Transparency reports often do not go into detail about what makes a user identifiable (e.g. do they just hold an email address or their full name?)

Studies on the relationships between anonymity and abuse

We have explored in detail a range of existing studies on the links between anonymity and abuse and have summarised the key studies and findings in the table below.

Study	Sample / Method	Key findings	Limitations
Meta analysis			

The effect of anonymity on conformity to group norms in online contexts A meta-analysis Guanxiong, H. & Li, K. 2016	Meta-analysis of 13 journal articles regarding anonymity and conformity to group norms	This research investigated the tension between two theories of how the self relates to groups. The *deindividuation model* argues that the 'loss of self' experienced from anonymity increases the likelihood of behaviour that betrays group norms. The *social identity model* argues that group membership has more influence on behaviour and the 'loss of self' increases conformity to group norms. The results support the latter: when anonymous, users were *more* likely to conform to group norms. This would suggest that the culture of online spaces has more impact on (abusive) behaviours than anonymity itself	The sample size of this meta-analysis was limited. A further limitation is related to the inherent weakness of meta-analysis in selecting publications for review, and the bias against publications/studies with nonsignificant findings. Only studies with significant findings were accepted and far fewer than have been published on the topic. Effect sizes may therefore be overestimated compared to population effect sizes
Theory			
Online Disinhibition Effect Suler, J. 2004	Explores six factors that interact with each other in creating the online disinhibition effect: dissociative anonymity, invisibility, asynchronicity, solipsistic introjection, dissociative imagination and minimization of authority	Users of the internet lose inhibition. This can work in two ways: benign disinhibition (unusual openness, vulnerability, generosity) and toxic disinhibition (rude language, hatred, harsh criticisms). Anonymity is a key factor that influences online disinhibition	Although this article provides useful analysis for theoretical frameworks, many other authors in this table dispute Suler's findings The article includes only 12 references
Anonymity, Pseudonymity, and Deliberation: Why Not Everything Should Be Connected Moore, A. 2018	Applies political theory to the issue of anonymity; analyses anonymity in terms of "two dimensions of identity disclosure": durability and connectedness	It is important to separate two aspects of anonymity that impact behaviour – durability (having a lasting identity and reputation) and connectedness (your behaviour in one space impacting your reputation in another, for example if you made racist comments online you could get fired from your job) Suggests that platforms should enable durable identities that makes it harder for trolls and abusive users to act with impunity, but without demanding connectedness to offline identities	This article provides useful theoretical frameworks for research, but is not itself evidence based

		Adds nuance to the anonymity/ disinhibition debate: highlights the possibility of creating online spaces where users can be disconnected from their offline selves, whilst simultaneously held to account/needing to uphold their reputation	
Experiment/controlled study			
Effect of anonymity, invisibility and lack of eye contact on online disinhibition Lapidot-Lefler, N. & Barak, A. 2012	Used an experimental design with 142 participants to study the impact of three typical online communication factors on flaming behaviours	Separates factors that influence online disinhibition: anonymity, invisibility and lack of eye contact. Finds that lack of eye contact was the chief contributor to toxic disinhibition. This study implies that anonymity's role in abusive behaviour has been over-emphasised and other experimental studies have defined anonymity too broadly by conflating it with invisibility	The design of this research was experimental, with limited external validity Although this study neutralized the gender variable (due to its impact on eye contact/social interactions) it did not include gender matching or account for other individual differences which may have influenced the results
Changing Deliberative Norms on News Organizations' Facebook Sites Stroud, N. J., Scacco, J. M., Muddiman, A. & Curry, A. L. 2014	Analysed comment section norms on 70 political posts from a local television station by assessing the impact of three different test subjects engaging in the comment sections - A) an unidentified staff member B) a recognizable political reporter and C) no one	Having an identified reporter engage with commenters affected the deliberative tone of the comments: there were lower levels of incivility and a greater use of evidence. Indicates that engagement from moderators can shape norms of interaction in comment threads (this aligns with literature on the role of opinion leaders in shaping group norms and research that argues the culture/norms of online platforms are more decisive when it comes to abuse than anonymity)	The results of this study are not generalisable – it uses data from a single Facebook page and analyses the effects of just one reporter
Preventing Online Harassment and Increasing Group Participation Through Social Norms in 2,190 science discussions	R/science thread – 13.5m subscribers Team of 1200 volunteers reviewed 62,457 comments over 29 days and tested the effect of posting rules on the behaviour of first-time commenters. Commenters included anonymous and identifiable accounts	Posting the rules causes an 8.4-percentage-point increase in the chance that a newcomer's comment will be allowed to remain by moderators on average This study indicates that other factors beyond anonymity can influence individual online behaviours. Interventions by moderators can influence social norms,	These results are not generalisable – they are based on a single Reddit thread and the study does not robustly account for other variables which may impact online behaviour, such as platform culture or design of the software

Nathan Mathias, J. 2019		i.e. posting rules on discussion groups and evidence. However, it was only on one thread (R/science) so behaviour might differ in others	
Survey			
Cyber Bullying Behaviors, Anonymity, and General Strain Theory: A Study of Undergraduate Students at a South Eastern University in the United States Lee, G. & Sanchez, M. 2018	Survey questionnaires sent to 15 undergraduate classes (406 students) at South Eastern University in the US	Strain caused by academic shortcomings (being threatened with or actually losing a scholarship) led to more cyberbullying behaviours Anonymity was negatively associated with cyberbullying and cybervictimisation – a higher level of internet anonymity was shown to reduce cybervictimisation	The sample is small and covers just a single cohort of the public: university students. A larger, more diverse sample which was representative of the population would be more reliable
Three-quarters of those experiencing online abuse say it comes from anonymous accounts Compassion in Politics. 2021	2003 respondents were asked "To what extent would you support or oppose government action to reduce the number of anonymous accounts on social media platforms?"	47% - strongly support 26% - somewhat support	The survey question does not offer a more detailed definition of what is meant by 'anonymous' nor what kind of 'government action' would be enacted The sample is small and covers just a single cohort of the pubic, university students. A larger, more diverse sample which was representative of the population would be more reliable

['I did it for the LULZ': How the dark personality predicts online disinhibition and aggressive online behaviour in adolescence](#) Kurek, A., Jose, P. E. & Stuart, J. 2019	709 high school students completed a survey assessing personality and technology behaviours. This survey was designed to investigate whether the three 'dark' personality traits of narcissism, psychopathy, and sadism would predict false self-perceptions, and in sequence, online disinhibition and aggressive online behaviour	All three 'dark' personality traits were positively associated with online disinhibition and cyber aggression This study suggests that other variables aside from anonymity have an impact on online abuse. It challenges the idea that anonymity *changes* behaviour – users' offline personalities are a better determinant of abusiveness	The sample is small and covers just a single cohort of the public: high school students. A larger, more diverse sample which was representative of the population would be more reliable
Data scraping/ social listening/ discourse analysis/ content analysis			
[Digital Social Norm Enforcement: Online Firestorms in Social Media](#) Rost, K., Stahel, L. & Frey, B. 2016	Analysed all comments on online petitions published on the German social media platform www.openpetition.de over the course of three years. The final dataset included 532,197 comments on 1,612 online petitions	Only 29.2% of all commenters prefer to remain anonymous Both the random-effects and fixed-effects models show that more online aggression is obtained by non-anonymous commenters and not by anonymous commenters Suggests that aggressive commentators in a socio-political setting gain more status and recognition from being identifiable	The results are not widely generalisable; the sample is limited as it is drawn from a single online setting where the user base is narrow. The platform also includes unique motives for remaining identifiable which in consequence disallows inferences to be made between the impact of anonymity and abuse/aggression which can be broadly applied. Also, comments were checked qualitatively, and not all comments on the platform were analysed
[Virtuous or Vitriolic: The effect of anonymity on civility in online newspaper reader comment boards](#) Santana, A. D. 2013	An examination of user comments on US newspaper sites that allow anonymity (N=450) and the user comments on newspapers that do not (N=450), which compares the level of civility in both. This focused on articles on the topic of immigration	Anonymous commenters were significantly more likely to register their opinion with an uncivil comment than non-anonymous commenters. Just over 53 percent of the anonymous comments were uncivil, while 28.7 percent of the non-anonymous comments were uncivil Nearly 15 percent of the anonymous comments were	The newspapers included in the study all have different community guidelines, user bases and article styles, which the study was unable to control for

		civil, while 44 percent of the non-anonymous comments were civil. The remainder in both categories were either neither/nor or unclear. Concludes that removing anonymity has a significant impact on reducing the levels of incivility in newspaper comments	
'Anonymity, Membership-Length and Postage Frequency as Predictors of Extremist Language and Behaviour among Twitter Users' Carter, P. & Sutch, H. 2019	A total of 205 Twitter accounts and 102,290 tweets were examined Corpus linguistic analysis (CLA) and content analysis (CA) were conducted to assess Twitter users' extremism in relation to anonymity, membership length and postage frequency. The former looked at extreme words associated with Islam and the latter looked at four types of extremist behaviour (extreme pro-social, extreme anti-social, extreme anti-social prejudicial biases and extreme radical behaviours)	Twitter accounts who have high anonymity and where the number of identifiable items is low were significantly more associated with extreme words As well as users with high levels of anonymity being more likely to post extreme words, other factors also played a role. Shorter membership and a lower postage frequency also increased the likelihood of extreme words being posted	This research only explores extremism on Twitter and not other social media platforms. Given the variety in design, culture and community guidelines across platforms these findings, particularly around the impact of anonymity, are not generalisable
'Meaner on Mobile: Incivility and Impoliteness in Communicating Contentious Politics on Sociotechnical Networks' Groshek, J. & Cutino, C. 2016	Public content from Twitter was collected through the Boston University Twitter Collection and Analysis Toolkit (BU-TCAT) on an ongoing basis. BU-TCAT had archived over 375 million tweets at the time of publication	The research found that mobile communication is both more uncivil and impolite than fixed web messages	The study focused on identifying impoliteness, including sarcasm and hyperbole, rather than abuse. Intensity of hostility was not taken into account

The links between anonymity and abuse

The studies in the table above provide mixed conclusions on the link between anonymity and abuse. Many of the studies are limited in their focus, with differing definitions of abusive behaviour and some only exploring a single platform.

Several studies refer to John Suler's theory of the disinhibition effect, which suggests that online communication causes users to lose inhibition and act in ways that they would not offline. This can work in two ways: benign disinhibition (unusual openness, vulnerability, generosity) and toxic disinhibition (rude

language, hatred, harsh criticisms).[127] The principle is that when a person's online persona is not traceable back to their offline self, people feel freer to act differently.

Within the more recent studies, some have corroborated this theory. A 2019 study utilising a data scraping methodology on Twitter which analysed over 100,000 tweets for extreme words found a much higher proportion of extreme language from accounts that were anonymous.[128] In this study a user's level of anonymity was based on 'identifiable items' on their Twitter account, such as full name or partial name, identifiable profile picture, specific location, and any additional links to social media profiles or personal information. This represents a more thorough assessment of anonymity than other studies which focus on real names only, or profiles which are identifiable in some way to the platform itself.

Similarly, a smaller scale study of newspaper website comment sections found a significantly increased number of comments were uncivil on websites where anonymity was permitted, compared to on those where it was not.[129]

However, studies also exist which have concluded that anonymity is not a driving factor behind abusive or uncivil posts. For example, a controlled experiment of 142 participants explored the three factors of online communication that are understood to trigger the disinhibition effect: anonymity, invisibility and lack of eye contact.[130] It found that lack of eye contact was the chief contributor to toxic disinhibition. This suggests that simply not being face to face with other people has the biggest impact on people feeling more able to be abusive online, with further levels of anonymity having less of an impact.

Other studies exploring online behaviours have similarly found that the link between anonymity and abuse is not so conclusive. A questionnaire study, completed by 406 US university students, found that anonymity was negatively associated with cyberbullying behaviours—it was observed that when anonymity is high, students are less likely to engage in cyberbullying perpetration.[131] A large-scale data scraping study by Rost, Stahel and Frey, which analysed all comments on online petitions published on a socio-political German social media platform, also found that more online aggression was posted by non-anonymous commenters.[132]

What other factors contribute to online abuse / incivility?

The study by Frey et al. of posts on a German political platform www.openpetition.de indicates that other users on the platform, and the culture of a platform, have a significant impact on prevalence and types of abusive content, and whether or not anonymity plays a role. The study explored the impact of social norm theory, where 'online firestorms' (collective online aggression directed towards actors of public interest) can enforce social norms by expressing public disapproval:

"In online firestorms, large amounts of critique, insulting comments, and swearwords against actors of public interest are propagated in social media within hours. This article begins the investigation on this rather new phenomenon by introducing a novel view on online aggression in social media. Relying on social norm theory, we proposed and demonstrated that one major motivation for online aggression in social media is the enforcement of social norms".

Within the culture of www.openpetition.de as a political social media platform, anonymity would lead to a lack of credibility, meaning users had limited incentives to post anonymously, and posting negative comments towards public figures brought status to those involved. This suggests that anonymity fuels abuse only in certain instances, and is in fact very dependent on the culture and purpose of an online platform or website.

---

[127] Suler, J. 2004. The Online Disinhibition Effect. *CyberPsychology & Behaviour* 7(3) 321-326.

[128] Carter, P., Sutch, H. 2019. Anonymity, Membership-Length and Postage Frequency as Predictors of Extremist Language and Behaviour among Twitter Users. *International Journal of Cyber Criminology,* 13(2): 439-459.

[129] Santana, A., 2013. Virtuous or Vitriolic. *Journalism Practice*, 8(1), pp.18-33.

[130] Lapidot-Lefler, N. and Barak, A., 2012. Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior*, 28(2), pp.434-443.

[131] Lee, G. & Sanchez, M.. (2018). Cyber Bullying behaviors, anonymity, and General Strain Theory: A study of undergraduate students at a South Eastern University in the United States. *International Journal of Cyber Criminology*, 12, pp.84-96.

[132] Rost, K., Stahel, L. and Frey, B., 2016. Digital Social Norm Enforcement: Online Firestorms in Social Media. *PLOS ONE*, 11(6), p.e0155923.

What could platforms be doing around anonymity and abuse?

An experiment in the r/science community[133] on Reddit[134], which at the time had 13.5m subscribers, asked moderators to post the community rules in a 'sticky' comment at the top of some threads.[134] This correlated with newcomer comments in discussions complying with community rules at higher rates: *"Without posting the rules, a first-time commenter in a discussion about a peer-reviewed article has a 52.5% chance of complying with community norms. Posting the rules causes an 8.4-percentage-point increase in the chance that a newcomer's comment will be allowed to remain by moderators on average in r/science."* The experiment indicated that interventions by moderators can influence social norms and affect online behaviour, and adds to the argument that the culture of a platform has a large impact on the amount of abuse present.

A platform like Reddit, therefore, with individual communities that are moderated by users, appears to have an increased ability to influence their culture through highly visible community guidelines and the strong presence of individuals there to enforce them. It would likely be more difficult for platforms that operate with less of a community-based model to replicate this (such as Instagram or Twitter). However, there will likely still be a number of opportunities for them to influence culture, and to increase the visibility, impact and respect for their guidelines.

Alfred Moore argues that it is important to separate two aspects of anonymity that are associated with abusive behaviour.[135] The first is a lack of durability: the idea that a user can create temporary anonymous accounts which free them from the reputational pressures of a lasting persona, i.e. creating 'burner' accounts to abuse other users. The second is a lack of connectedness: the idea that a user's actions in an online space are not traceable to them offline – they can act differently in offline and online spaces without accountability.

Enforcing connectedness would disable the positive aspects of the disinhibition effect, i.e. unusual openness, vulnerability and altruism online and, as outlined earlier, could negatively impact some individuals who need to remain anonymous online for reasons such as personal safety. However, by enforcing durability, Moore argues that it would be harder for trolls and abusive users to act with impunity, as in practice this would mean platforms reducing users' ability to create multiple accounts, and ensuring their actions have a long-term impact on their ability to post or interact with others.

Overall, the above studies suggest that platforms could be:

- Aiming to improve or change the platform culture to make abuse less appealing to users – ensuring it reduces status, views etc. rather than increasing them.
- Stopping users from creating multiple accounts which means there are no consequences or 'durability'. Currently if a user is blocked on one account they are often easily able to create another and carry on with their behaviour

---

[133] Reddit is a large community made up of thousands of smaller 'communities'. These smaller, sub-communities within Reddit are also known as "subreddits" and are created and moderated by users.

[134] Matias, J., 2019. Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. *Proceedings of the National Academy of Sciences*, 116(20), pp.9785-9789.

[135] Moore, A., 2017. Anonymity, Pseudonymity, and Deliberation: Why Not Everything Should Be Connected. *Journal of Political Philosophy*, 26(2), pp.169-192.

Existing policy recommendations

Various policy recommendations to tackle the issue of anonymous abuse have been proposed by different organisations, particularly in reference to the Online Safety Bill. Most recommendations acknowledge the benefits of anonymity online, and therefore do not advocate for an outright ban on anonymity or pseudonymity. Specific proposals are summarised in the table below, but broadly include:

- Explicitly including anonymity as a risk factor for harm (i.e., online abuse) in proposed risk assessments for social media companies
- Increasing friction into the process of creating social media accounts
- Ensuring law enforcement can access sufficient information from social media platforms to tie online abuse to perpetrators where a crime has been committed
- Providing social media users with the option to verify their identity
- Providing social media users with the option to block interactions with unverified users, and enabling users to see whether or not a user is verified

The table below summarises the existing policy recommendations. It is important to note that there has been a lot of material published and submitted in the pre-legislative scrutiny phase of the Draft Online Safety Bill. What is included below crosses over with this to a large extent, but focuses in on recommendations around *anonymity* in particular.

Organisation	Publications	Overview	Key points
Carnegie UK	Publications include but are not limited to: 'The Draft Online Safety Bill: Carnegie UK Trust initial analysis', Carnegie UK Trust, 2021 'Submission to Petitions Committee Inquiry into Tackling Online Abuse', Carnegie UK Trust, 2021	Carnegie UK Trust is a not-for-profit organisation focused on improving wellbeing through a range of research, advocacy and community programmes They have been heavily involved in the pre-legislative scrutiny phase of the Online Safety Bill	• Carnegie UK's solution is to include anonymity as something that each platform must include in their risk assessments even if there is no one-size-fits-all solution
Clean up the Internet	Publications include but are not limited to: 'Time to take off their masks? How tackling the misuse of anonymity on social media would improve online discourse and reduce	Clean up the Internet is an, independent, UK-based organisation concerned about the degradation in online discourse and its implications for democracy	• They state "anonymous and unverified accounts would be an obvious candidate to be designated as a risk factor, given both the clear evidence of the role of anonymity in enabling harm, and the public expectation that it be tackled" • They also address "legitimate concerns about measures to restrict abuse of anonymity". Key points are: A) Ensure that verification systems are developed with due regard to accessibility, diversity, and inclusion

	abuse and misinformation', Clean up the Internet, 2020 'Written evidence submitted by Clean up the internet', Clean up the internet, 2021	They focused on links between anonymity and abuse online. They have commissioned some research in the form of a YouGov poll which includes self-report evidence from members of the public around online anonymity and abuse	B) Place strict limits on the use of data collected for purposes of verification C) Make verification optional ● They suggest that anonymity has been identified as a key factor in the online disinhibition effect **Makes 3 policy proposals:** 1) Offer all users the choice to verify 2) Offer all users an option of choosing whether or not unverified users are able to interact with them 3) Underpin these two requirements by making social media companies who fail to implement measures to mitigate the negative effects of anonymity, such as those proposed above, legally liable for content produced by anonymous and unverified users
HOPE Not Hate	Publications include but are not limited to: 'Written evidence [to the Pre-legislative Scrutiny Committee for the Draft Online Safety Bill] submitted by HOPE Not Hate', (UK Parliament, 2021)	HOPE Not Hate are an anti-fascist organisation who monitor far right extremism online and offline	● Emphasise the importance of some level of anonymity to protect vulnerable groups. But call for more clarity around the OSB's definition of "Democratically Important" content, as this may come into conflict with harmful or hateful content ● Want measures to introduce friction into the process of creating and removing accounts, requiring accounts to build up evidence of rules adherence and compliance before accessing full functionality of a platform ● Suggest one way to maintain anonymity but reduce harm is by giving users more control over content they see. E.g. customisation of their own networks and limiting interactions with anonymous accounts, and making ID verification optional (similar to Clean up the Internet's suggestion) ● Emphasises any legislation relating to anonymity needs to focus on how platforms are built, designed and run. Including: **1)** Stopping algorithms promoting harmful and divisive content **2)** Encouraging anonymous but stable identities online
Antisemitism Policy Trust	Publications include but are not limited to: 'Regulating Online Harms: Tackling Anonymous Hate', Antisemitism	Explores the impact of online anonymity in relation to racism, extremism and discriminatory bullying on the basis of ethnicity or religion.	● States that since the emergence of COVID-19 there was a rise in antisemitic conspiracy theories and the use of antisemitic rhetoric and stereotypes with reference to the pandemic ● Argues that limiting online anonymity will provide victims of abuse with greater confidence and control and suggests that it would reduce online abuse

	Policy Trust, 2021	Argues that online anonymity is harmful enough to merit regulation	• Argues that platforms should determine the level of anonymity they allow their users to have but must follow a statutory duty of care to their users • Suggests it should be up to the platform to determine the degree of anonymity it wishes to give to users. But platforms should stipulate in terms and conditions that anonymous users engaging in hate speech and abusive behaviour will be banned and could have their identity revealed to law enforcement. More about underlining police powers than removing anonymity. • Supports verifying user identities (with electronic identification) before allowing people to sign up
Stonewall	Publications include but are not limited to: 'Stonewall – written evidence: House of Lord Communications and Digital Committee inquiry into Freedom of Expression Online', UK Parliament, 2021	Stonewall, Europe's largest organisation campaigning for LGBT+ equality – submitted evidence for the pre-legislative scrutiny phase for the Draft OSB	• Highlights potential unintended consequences of legislating around anonymity on marginalised groups • "Any approach that requires individuals to verify their identity to a third party with a birth certificate, or other forms of personal ID, raises significant privacy and security concerns" • Provides no concrete recommendations beyond recommending a cautious approach before legislating around anonymity, as it could limit freedom of expression
International Peace Institute	Publications include but are not limited to: 'At the Nexus of Participation and Protection: Protection-Related Barriers to Women's Participation in Northern Ireland', International Peace Institute, 2021	The International Peace Institute (IPI) is an independent, international not-for-profit think tank dedicated to managing risk and building resilience to promote peace, security, and sustainable development Report based on 25 interviews with women in Northern Ireland describing the landscape of women's exclusion from participation in	• Suggests online abuse in Northern Ireland has an extra layer shaped by the history of the conflict which includes sectarianism, conflict-related violent masculinities and a unique misogyny affected by both • Argues anonymity has a role in encouraging online abuse of this type • Argues that anonymity on platforms makes it harder to hold the attackers or the publishers to account for abuse • Anecdotal evidence from the interviews conducted suggests that online abuse greatly impacts women's private lives and often takes place when they talk publicly about progressive ideas which do not align with sectarian perspectives

		peace-building efforts in Northern Ireland	
Glitch	Publications include but are not limited to: 'The Ripple Effect: COVID-19 and the epidemic of online abuse', Glitch, 2020	Glitch is a charity with a mission to end online abuse and champion digital citizenship across all online users Report is about online abuse against women during the COVID-19 pandemic, based on a survey taken in 2020 aimed at women and non-binary individuals – had 484 responses	• Suggests that government should adopt a public health approach tackling online abuse • Specifically in relation to anonymised data, they recommend that tech companies need to provide greater transparency about their content moderation including allowing trusted research institutions to access anonymised data about content removals and complaints submitted to platforms, including the type of action taken, the time it takes to review reported content and increased transparency around appeals processes
Demos	Publications include but are not limited to: 'Is anonymity the best target in the fight against the trolls?' Demos, 2021	Thinktank blog post recommending that anonymity and anonymous abuse should not be conflated, and require different policy approaches	• Argues anonymity and anonymous abuse should be considered separately when being addressed by policy, and tighter restrictions on anonymity should not be the answer to reducing anonymous abuse • Suggests that reducing anonymity online will only change how online abuse is responded to, and may not solve it at its root • Believes greater identification is not prevention focused • Referring to social norm theory, the piece argues that efforts should be made to encourage friendlier, safer online spaces where abuse is not tolerated as part of the culture of the platform

International approaches to dealing with anonymity and abuse

So far, few countries have successfully implemented measures which seek to reduce anonymity in order to tackle abusive behaviours online. Countries which have existing measures in place have included nations with reduced freedoms for their citizens and a history of censoring content online, such as China.[136]

In 2007 South Korea put in place legislation which required the largest Korean-based platforms to collect ID information for people posting on their websites.[137] However, this was overturned in 2012 as it was found to be a disproportionate response to the issue, resulting in data breaches and with limited evidence that it had any impact on online abuse.

More countries are currently exploring ways to make platforms more responsible for identifying their users. For example, Australia is planning to make social media companies collect and store identifiable information.[138] France is also exploring the idea of linking newly released Digital ID to social media platforms – which would require users to upload identification upon registration to the platform.[139]

All of these example measures still allow users to create profiles which do not require their real name—therefore remaining anonymous to other users. However, the platforms are (or will be) required to collect and store identifiable information on users which will increase accountability for their actions online.

The table below outlines different countries' approaches to anonymity.

Country	Sources	Approach	Outcomes
Germany	Arthur Piper, International Bar Association. 'Time to tackle internet anonymity'	• German courts ruled that a clause in Facebook's user terms broke the law because people had to agree to use their real names when signing in on the platform – despite the fact that they had no clear idea of how those details would be used • The German Telemedia Act maintains that it must be possible to use any telemedia service as long as it is technically possible and feasible for the provider	• Upholds the idea that everyone should have access and be able to use online services, without having to share their real identity • Has reduced the ability for platforms to identify users as part of maintaining data privacy
Australia	Federal Register of Legislation, Online Safety Act 2021. 'An Act relating to online safety for Australians, and for other purposes'	• The government, through their Online Safety Act, has imbued the e-safety commissioner with powers to force social media companies to hand over the personal information of people the	• The law does not currently make any changes to how users present themselves on social media, but will mean the details that are held on them by platforms

---

[136] Liao, S., 2017. *China is forcing internet companies to end online anonymity.* [online] The Verge.

[137] Taylor, J., 2021. *Twitter says any move by Australia to ban anonymous accounts would not reduce abuse.* [online] The Guardian.

[138] Healey, B., 2021. *Australia's new Online Safety Act is starting to take shape, with harsh penalties levelled at trolls to include fines of up to $111,000.* [online] The Chainsaw.

[139] Mascellino, A., 2021. *Social media accounts could soon require digital ID in France,* UK | *Biometric Update.* [online] Biometric Update.

	Bianca Healey, Business Insider, 2021. 'Australia's new Online Safety Act is starting to take shape, with harsh penalties levelled at trolls to include fines of up to $111,000' Josh Taylor, the Guardian, 2021. 'Twitter says any move by Australia to ban anonymous accounts would not reduce abuse'	commissioner alleges are engaging in bullying behaviour online. These new powers will come into effect at the start of 2022	can be handed over to authorities
China	Shannon Liao, the Verge, 2017. 'China is forcing internet companies to end online anonymity'.	● As of 2017, social media companies were told to ensure all users were identifiable. Users can still use pseudonyms, but they must be tied to their real identities ● Internet companies and service providers are being made responsible for ensuring users stay fully identified ● Companies and service providers are also required to report any illegal content they see on any platform to the government	● The end of online anonymity comes alongside a range of other restrictions for online users in China ● This will likely greatly reduce any instances of anonymous abuse and illegal activity, but freedom of speech and many types of content are also heavily censored
South Korea	Josh Taylor, the Guardian, 2021. 'Twitter says any move by Australia to ban anonymous accounts would not reduce abuse'	● A law was in place between 2004 and 2012 which required the largest Korean-based websites to collect ID information for people posting on their websites in a bid to reduce cyberbullying online ● This was called the 'Real Name Verification Law'	● When it was overturned, the court found there was insufficient evidence to demonstrate that there had been a decrease in hateful comments, defamation or online bullying as a result of the policy ● A breach of 35 million South Koreans' national identification numbers was one issue that led to it being overturned
Japan	Freedom House, 2021. 'Freedom on the net 2021, Japan'	● In April 2021, a bill was unanimously passed which made it easier to identify users who allegedly slander people on the internet. The law is scheduled to take effect by the end of 2022. Under the amended law, an individual can request the court to disclose information about a sender	● There are concerns that disclosure requests will be misused as a means to suppress the transmission of information

		- who posted defamatory content - The court has up to six months to decide whether to order a content provider to disclose the sender's information, and can also order access providers to retain the sender's information during the proceedings	
Brazil	Renan Araujo and Alice Gaudiot, Oxford Human Rights Hub, 2020. 'Brazil's 'fake news' bill threatens to harm internet freedom and individual rights' Diogo Tulio dos Santos, Global Americans, 2021. 'Brazil, democracy, and the "fake news" bill'	- Brazil is considering legislation that would compel social media platforms to identify all users and their locations to prohibit 'unauthentic accounts' - A bill to 'fight fake news' was approved by the senate in June 2020. If passed it would oblige platforms to keep individual records that 'guarantee unequivocal identification' of users for at least six months	- Is being talked about by human rights organisations as a bill that would harm online freedoms - Facebook, Twitter, and Google also released a joint statement opposing the bill. They refer to it as a "project of mass collection of data from individuals, resulting in worsening digital exclusion and endangering the privacy and security of millions of citizens"
France	Alessandro Mascellino, Biometric Update, 2021. 'Social media accounts could soon require digital ID in France'	- The website of the French Senate has published a new bill proposing the creation of an independent supervisory authority in charge of collecting user identities when they register with online platforms - According to the new document, submitting a scan of their ID upon registration would make users aware that they can be identified quickly, and therefore serve as a deterrent against offending behaviours	France is already rolling out Digital ID cards, however platform verification has not yet been made into legislation
Canada	Antisemitism Policy Trust, 2020. 'Regulating Online Harms: Tackling anonymous hate'	- The Pan Canadian Trust Framework (PCTF) is a set of digital ID and authentication industry standards - The PCTF's principles include asking users to provide only the minimum amount of personal information, privacy enhancing tools such as the	Has been rolled out gradually across organisations and users

		'right to be forgotten', inclusion and transparency	

Conclusion

The links between abuse and anonymity are widely contested across the studies explored within this rapid evidence review. However, many studies are limited in their methodology or scope, with variations in definitions of abuse, limited sampling criteria or a lack of consideration for contextual information and severity of abuse.

Despite this, there are several large-scale studies which do suggest anonymity on platforms increases the likelihood of abusive, or uncivil, behaviours. However, removing anonymity is rarely suggested as the best solution to reducing abuse, as there are a range of other factors which must be taken into account when assessing the issue:

- Abuse still occurs from non-anonymous users, and there are studies which have found that – on certain platforms – anonymity does not lead to increases in abusive content, and in fact identifiable accounts are occasionally preferred for directing abuse at other individuals.
- Anonymity is highly valued by many. Individuals experiencing vulnerabilities often benefit from being able to remain anonymous online and feel that removing their ability to be anonymous may increase the abuse they personally experience, both online and offline.
- There is a lack of research on the impact of anonymous versus non-anonymous abuse on victims. Within this rapid review only one study was found which explored impact, where respondents reported mixed views on whether anonymous or non-anonymous abuse was more impactful.[93] This is an important line of inquiry as resulting harm is an essential factor in deciding what action needs to be taken. For example, if non-anonymous abuse is found to be less prevalent but significantly more harmful, this could have a large impact on the outcome of the debate.

Some studies have suggested other ways to reduce the link between anonymity and abuse, such as ensuring anonymous profiles are at a minimum identifiable to the platform, and changing the culture of an online space. According to many studies exploring social norm theory and anonymity, platform culture and the behaviours of other users have a large impact on behaviour, with some platforms or experiments attempting to harness this through changes in platform design and interventions by active and identifiable moderators.[134]

Internationally, some countries have taken steps to reduce anonymity online. Often measures are preferred which ensure users are identifiable to platforms rather than other users – meaning they can be personally held accountable for their actions online and referred to law enforcement if required.

This approach reduces the impact on individuals who might suffer from not being able to remain anonymous online. However, there is limited data on how effective the strategy of ensuring users are identifiable to the platform is in reducing abuse. In South Korea they found little evidence of impact and removed the policy.[137]

Overall, there appear to be approaches to limiting anonymity which are likely safe and acceptable for many users. However, until more is known about the definite links between abuse and anonymity, it is not possible to conclude whether this will have any positive effect. Similarly, even if much online abuse is currently anonymous, removing anonymity may just lead to an equivalent increase in non-anonymous online abuse.

To inform future action and policy, more research is needed which specifically explores the links between abuse and anonymity, the motivations of anonymous and non-anonymous users posting abuse, and the impact anonymous abuse has on users. The qualitative element of this research aims to add insight to this current gap in research.