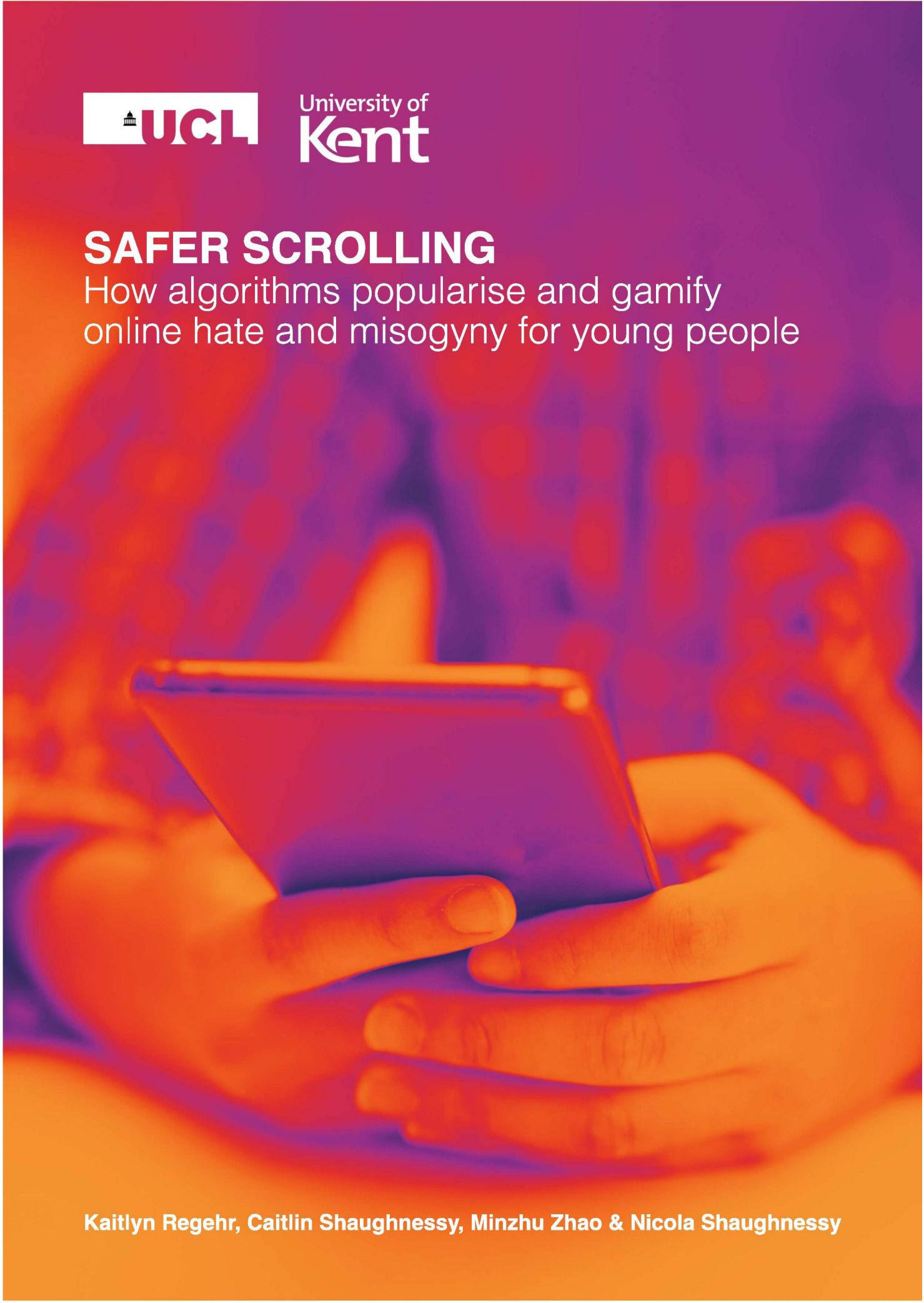
UCL University of Kent

SAFER SCROLLING

How algorithms popularise and gamify online hate and misogyny for young people



Kaitlyn Regehr, Caitlin Shaughnessy, Minzhu Zhao & Nicola Shaughnessy

ASCL000007_0001

ACKNOWLEDGEMENTS

This research is part of an Arts and Humanities Research Council (AHRC) funded project ("Understanding the Cel: Vulnerability, Violence and In(ter)vention", Principal Investigator: Dr Kaitlyn Regehr, Co-Investigator: Professor Nicola Shaughnessy) and additionally supported by the UCL Public Policy Rapid Response Fund. It has been conducted in partnership with University College London and University of Kent. The authors would like to express their extreme gratitude to all those who took part in the research, including the young people and school leaders who gave up their time to be interviewed. Further thanks go to the wider research team and experts who guided the project, in particular **Idil Cambazoglu**, for her support on data transcription, analysis and curation, **Alfie Turner,** for his support with participant recruitment and interviewing young people, and to **Margaret Mulholland** and the team of **Association of School and College Leaders (ASCL)** for their invaluable support and partnership throughout the research.

Research Team
Principal Investigator: Dr Kaitlyn Regehr
Project Manager and Postdoctoral Researcher: Dr Caitlin Shaughnessy
Co-Investigator: Professor Nicola Shaughnessy
Community Collaborator: Alfie Turner
Research Assistant: Minzhu Zhao
Research Assistant: Idil Cambazoglu

TABLE OF CONTENTS

Executive Summary 4

Introduction 6

Background 9

Methodology 12

Long form interviews with young people 12

TikTok Recommender Algorithms: archetype modelling 13

Expert consultation 15

Findings 17

1 Recommender algorithms and the gamification of radical ideology 17

2 Incel 2.0 and the normalisation of radical, toxic misogyny 24

3 Neurodiversity, existing vulnerabilities and mental health 27

4 Including young men in the conversation 31

5 Building critical digital literacy to support schools, parents and young people 32

Evidence based recommendations 34

References 36

EXECUTIVE SUMMARY

This report presents the findings from research conducted as part of an Arts and Humanities Research Council (AHRC) funded, interdisciplinary project, hosted at University College, London (UCL) in partnership with the University of Kent, which explores the popularisation of online hate speech and misogyny, and its risks to young people. It has been undertaken in collaboration with the Association of School and College Leaders (ASCL), who had identified concerns regarding young people's wellbeing and social media usage. This research has explored how social media platforms, and their algorithmic processes, are putting young people at risk through exposing them to increasingly harmful material, which is now presented as entertainment in young people's feeds. As a result, hateful ideologies and misogynistic tropes are becoming normalised in young people's behaviours both online and offline.

With this report, we highlight how the affordances of social media platforms, exemplified in this research by the recommender systems of TikTok, actively amplify and direct harmful content to young people. Through intensive online fieldwork, we employed an algorithm study which used an archetype modelling methodology (based on long-form interviews with young people) to analyse over 1000 videos on the social media platform TikTok. After only 5 days of TikTok usage, our research shows a **fourfold increase in the level of misogynistic content** being presented on the "For You" page of an individual's TikTok account. We found that the algorithm privileges more extreme material, and through increased usage, users are gradually exposed to more misogynistic ideologies which are presented and gamified through soft or humorous cultural forms.

Further, drawing from consultations with young people and school leaders, we have traced how these ideas become embedded within mainstream youth cultures, creating new challenges for schools. Social media's sophisticated technological affordances are providing a potent indoctrination effect, and this micro-dosing on highly toxic content is leading to the saturation of extremist misogynistic ideas among young people. Interviews with school leaders highlight the need for supports and strategies, and for coordinated messaging across constituencies, including parents, teachers, and technology companies. At present, schools and parents are concerned about what young people are accessing online, but have limited awareness of how harmful algorithmic processes function, of their impacts on young people and indeed, the wider community. Although for the purposes of this research we have explored how the TikTok recommender system operates, this research has wider relevance across multiple social media platforms, including Meta and YouTube. Indeed, it also has has wider implications for how parents and educators interact with and talk to young people about their social media habits.

Our research has highlighted that understanding and safeguarding technology use among young people requires a functional approach that examines the hybrid ecosystems of digital engagement, as well considering both the positive and negative consequences of different digital technologies. While young people express that online communities can make them feel understood, they are also aware of the potential for harm that online platforms can present but are unsure about how to navigate this harm. Further, adults are often unaware of how algorithmic processes can become harmful, or indeed of their own social media addictions, making parenting around these issues difficult. What is needed is investment in digital literacy – or a healthy digital diet - that equips and empowers young people with the critical digital skills to navigate their online worlds as they transition to adulthood.

Key Findings

1 **Harmful content is gamified and presented as entertainment through the algorithmic processes of social media platforms.**

Harmful content is presented as entertainment through the algorithmic processes of social media platforms (such as TikTok), which can amplify negative materials to young people and vulnerable groups. Similar to other studies, which have looked topics such as the alt right (Ribeiro et al., 2020), or self-harm material (Amnesty International, 2023), **our study found that after only 5 days of TikTok usage, there was a four-fold increase in the level of misogynistic content being presented on the "For You" page on TikTok.** In this way, toxic, hateful or misogynistic material is pushed to young people, exploiting adolescents' existing vulnerabilities. Boys who are suffering from poor mental health, bullying, or anxieties about their future are at heightened risk.

2 **As a result, ideologies, such as sexism and misogyny, are normalised amongst young people and seep into their everyday interactions.**

The proliferation of misogynistic ideas and language has now moved off screens and into school yards, where they are frequently enacted in mainstream youth culture. Young people increasingly exist within digital echo-chambers, which normalize this rhetoric. This is now seeping into everyday interactions between young people, with boys lacking awareness of its impact on their female peers.

3 **In response, it is essential to include boys as part of discussions regarding online misogyny, as hardline approaches can entrench negative ideologies and ways of thinking.**

Hard line approaches can have negative consequences, alienating young men. Instead, it is important to champion youth voice and include approaches that emphasise the role of peer-to-peer learning and positive male role models can promote an educative rather than punitive response to boys' behaviours.

4 **There is a need for a healthy digital diet approach to education to support young people, schools, parents and the community at large.**

Young people now exist in hybrid realities that are woven between their offline and online worlds. Schools face unprecedented challenges in this area, and outright bans on phones or social media are likely to be ineffective. Additionally, these approaches are not equipping young people to deal with these issues and content when they come across them. Instead, information regarding critical digital literacy – or a healthy digital diet approach – is needed to provide young people with key skills to recognize radicalisation and think critically about toxic online material. Further, adults are often unaware of how harmful algorithmic process function, or indeed their own social media addictions, making parenting around these issues difficult.

INTRODUCTION

In 2020, the online campaign 'Everyone's Invited' sparked a national conversation about the perpetuation of misogyny, harassment, sexual abuse, and sexual assault among young people in the UK. The campaign asserted that outdated sex education within schools was a major part of the problem, and that tackling rape culture required a cultural shift that recognized the pervasiveness of misogyny within education. The rapid report released by Ofsted in 2021 on sexual abuse in colleges and schools further highlighted the scale of the problem, with 9 in 10 girls reporting some level of sexist abuse or image-based sexual harassment (Ofsted, 2021). Research undertaken at UCL by members of this research team similarly found widespread digital sexual violence, where Image-based Sexual Harassment (IBSH) and Image-Based Sexual Abuse (IBSA) were reported by the majority of young women and girls (Ringrose et al., 2021). Arising from that report was a call to explore how gendered contexts like 'lad banter' were developing in online contexts. Recommendations were made to expand conversations around digital literacy with young people to develop 'trusting and honest relationship around online activity'.

These reports amplified growing concerns regarding impacts of digital engagement on adolescent development, with particular attention paid to young people's mental wellbeing and the paucity of research, in particular longitudinal research, that explores the links between the two. Of recent focus has been a move away from 'screen time' guidance, that is, the *amount* of screen time for adolescents, but instead a focus on the impacts of the *type* of screen time that young people are engaging with. The recent Online Safety Act in the UK is an indication of some progress to tackle the negative types of content that are frequently found online, but it doesn't sufficiently address the impact of algorithms on social media. Although this research focuses on one popular social media platform - TikTok - algorithms are used to similar effect across social media platforms. Indeed, reports such as those by Ribeiro et al.'s (2020), Reset Australia (2022) and Amnesty International (2023) have highlighted how the affordances of social media platforms, such as recommender systems of YouTube (where content is presented to users not only based on who they follow, but what the algorithm identifies as their interests) actively amplify and direct harmful content, in particular those associated with the alt right (Ribeiro et al., 2020), misogynist and manosphere content (Reset Australia, 2022), and self-harm material (Amnesty International, 2023). In a worldwide first, a recent ruling by the UK coroner in October 2022 further identified Instagram as 'likely contributing' to a young person's death, due to the high level of self-harm material recommended on her Instagram feed in the months before she died. Online echo chambers and the rapid rise and circulation of electoral fake news, mis/disinformation, and polarisation towards political extremism on the internet has had particular impacts for young people (Floridi, 2016; Krasodomski-Jones, 2017; Lazer et al., 2018; Lewis-Kraus, 2022). Most recently, misogynist influencers, who have leveraged their fame to promote polarized far-right extremism, have utilized the phenomenon of echo chambers to bring (gendered) hate to prominence in other dimensions of public discourse, such as boys' misogynist behaviours within schools and educational settings (Das, 2022).

This research began by interviewing young people engaging with and producing radical online content on radical Discord servers, in particular content related to online misogyny and Incel (involuntary celibate) content. This led us to trace how the same discourses, memes and tropes were now becoming part of mainstream material on social media. Content that was once only found in dark corners of the internet, relegated to subversive platforms like 8chan, or specific "alt" servers on Discord, was now algorithmically offered and circulated on popular teen platforms like TikTok. On these platforms, users are gradually exposed to more and more misogynistic ideologies which are presented and gamified through "soft" cultural forms, masking the toxic and violent misogyny at its roots. Our subsequent analysis of TikTok videos, and the processes by which particular themes and materials are pushed to vulnerable young people, highlight how high usage of social media platforms are leading to deeply unhealthy shifts in the way young people think and interact with their peers. As our interviews with school leaders have highlighted, despite significant positive changes in PSHE curriculums, there remain concerns around the manifestation of misogynistic behaviours between peers and towards female members of staff. Boys and young men are increasingly exposed to violent, toxic misogyny online and schools are facing new challenges in equipping young people with tools to engage more critically with the digital space and content that they are exposed to online.

Schools have further become the battleground for future digital literacy education and policy. Recent policy announcements regarding the banning of smartphones in schools, and new legislation on new age limits on social media for young people, bear little relevance to contemporary education challenges, which seek to support young people with the key digital skills needed to succeed in the future. As this report highlights, the exposure to and amplification of violent misogyny on social media, and its manifestations in adolescent interactions, are a symptom of a wider problem regarding digital literacy. In order to properly support and prepare young people for the future, wider conversations about how critical digital literacy can be embedded across into young people's skill sets are needed.

GLOSSARY OF TERMS

Term	Description
Incel	Short for *involuntary celibates,* incel refers to a person, usually a man, who express resentment, rage, and hostility towards women for denying them relationships and sex. The term originally emerged in online subcultures, with men fantasizing and sharing dark humour about violence against women.
Discord	A social media platform where users can communicate using audio calls, video calls, text messaging, as well as share media and files. Users can communicate privately to each other, or in virtual communities called "servers" and "channels".
TikTok	A social media app based on short-form videos. Users can submit videos, which can range from 1 second to up to ten minutes, often shot on smart-phones or webcams. On the app, users swipe or scroll up to navigate through videos on their "For You" page.
Algorithm	A computer-generated set of rules and data that control how content (such as videos, users, or posts) are filtered and recommended to users on the platform.
Hegemonic Masculinity	The ideology that men should be dominant in society at the expense of the subordination of women.
Red Pill	Concept of 'Red Pill' originated on Incel forums, and are a self-identified group that take the view that society is dominated by women but individual attempts such as learning "game" or improving your appearance can lead you to achieve a sexual relationship with women.
Black Pill	Concept of 'Black Pill' originated on Incel forums. Unlike Red Pills, Black Pillers believe that looks and relationships are genetically pre-determined, and that there are wider societal barriers to relationships. Black Pillers believe that women only care about men's physical looks, and that no individual actions will ever improve their chances of having a sexual relationship.
Misogyny	Ingrained prejudice, negative opinions, or contempt of women.
Manosphere	Umbrella term that refers to content, communities, forums, and websites that promote masculinity, men's rights and oppose women's rights and feminist movements.
Looksmaxing	Physical improvement of someone's body and appearance, such as through plastic surgery, diet, exercise or skin care.
Neurodiversity	Neurodiversity describes how people understand the world in different ways and the multitude ways that brains can process information. Often used as an umbrella term to describe a number of different conditions including Autism, ADHD, Dyslexia or Learning Difficulties.

BACKGROUND

Since the advent of the iPhone in 2007, most young people have grown up interacting with digital products from birth. 91% of young people aged between 12 and 15 years in the United Kingdom (UK) own a smartphone (Ofcom, 2021). Some estimates put screen time among young people between 5 and 6 hours a day (Twenge, Martin & Spitzberg, 2019), with others suggesting it is up to 9 hours a day (AACAP, 2020). For young people, the majority of this screen time involves using social media and the way they interact with online material differs from previous generations. Young people today may be considered *social natives* – who largely grew up in the world of the social, participatory web, and 'differ meaningfully from *digital natives* (25–34s) – who largely grew up in the information age but before the rise of social networks – when it comes to news access, formats, and attitudes' (Reuters, 2022, p. 41). As a report by Reuters (2022) found, young people are moving towards Instagram, TikTok and YouTube not just as a way to interact socially, but as a way to consume and access information: 'Use of TikTok for news has increased *fivefold* among 18–24s across all markets over just three years, from 3% in 2020 to 15% in 2022, while YouTube is increasingly popular among young people in Eastern Europe, Asia-Pacific, and Latin America' (Reuters, 2022, p. 41). The affordances of these platforms, and in particular the risks that they present in terms of disinformation and radicalisation has been highlighted by many (Floridi, 2016; Jackson, 2017; Lazer et al., 2018; Lewis-Kraus, 2022), yet research is often limited by social media companies themselves, who obfuscate and limit access to researchers seeking to understand the impact of their recommender algorithms (Amnesty International, 2023).

Wider studies in this area have also highlighted how high dosages of hateful content, including extreme, violent views, are becoming normalised for young people through their usage of social media. For example, Matamoros-Fernandez and colleagues (2021) found that YouTube's algorithm boosts authoritative outlets and recommends misleading videos regardless of the actual contextual background of the content (e.g., Jordan Peterson and other misogynist content was being recommended as 'winners' for the 'feminism' key word, pp.240-242). A 2022 report published by Reset Australia on extremist misogynist algorithms on YouTube showed similar results. By using an auditing methodology based on setting up '10 avatar accounts on YouTube aged between 13 and 20, (Reset Australia, 2022, p.6), the study demonstrated how algorithms readily feed misogynist and anti-feminist content of Jordan Peterson into their users' featured clips within a short period of time. By exploring patterns in YouTube recommendation algorithms to boost and prioritize homogenous Manosphere content to their users, the study confirms the link between liking Peterson's videos and exposure to further misogynist content.

Online Misogyny and Incel 1.0

There were also increasing concerns regarding the growth of Incel cultures, (online communities of individuals who feel rejected by women), and how technology was facilitating processes of indoctrination that were transforming into misogynistic extremism (Regehr, 2020). Incel has referred to a population of young people who feel left out of romantic relationships, and often, society more generally, who take to the digital space to voice feelings of frustration, anger and at times, a desire for revenge. Typically, these individuals feel like outsiders, and struggle with social anxiety (Daly & Reed, 2022; Regehr, 2020; Speckhard & Ellenberg, 2022). Incel communities have traditionally embraced 'nerdom', or what Massanari (2017) has termed 'geek masculinity'. In addition, they also often take on the language of the oppressed (Ging, 2019) and typically focus on women as their source of unhappiness. In the end, they may be indoctrinated into misogynistic extremism.

In a survey carried out by members of the community in 2019 on i.imgur.com, all of 667 respondents were male, and 64% were between the ages of 18 and 24. Among respondents, 70.8% suggested that they had been bullied in childhood and 94.7% believe they had missed out on developmental milestones. With respect to mental health concerns, 67.5% said they experienced 'long lasting depression'; 74.1% identified that they experienced anxiety, stress, or emotional distress 'in a constant manner'; 76.5% indicated that they had seriously considered suicide (League, 2020).

Concerns have been expressed regarding the risks presented by the community of young men who self-identify as Incels (Hoffman, 2019). Engagement with the Incel community has been associated with both self-harm and serious incidents of mass violence, particularly in the United States and Canada, and an attempted incident in Scotland (Horne, 2021). Incel related attacks have taken the form of shootings, stabbings, and vehicular homicides, carried out on university campuses, in schools, and city centres. Though the number of cases that scholars have definitively linked to Incel violence varies from seven (Tomkinson et al., 2020), to 15 cases (Hoffman et al., 2020; Wood et al., 2022), the threat has been clearly documented. Incel has been regarded as posing a serious terrorist threat: 'Incel violence arguably conforms to an emergent trend in terrorism with a more salient hate crime dimension that necessitates greater scrutiny and analysis- especially as it spreads to Europe and shows similarities to and has nascent connections with other terrorist movements, (Hoffman et al., 2020, p. 565). Further, as Diaz and Valji (2019) have argued, there is a clear connection between misogyny and physically violent extremism (Díaz & Valji, 2019). Digital misogyny may act as a gateway to or early warning sign of actualised violence.

Popularisation of Incel 2.0

What has emerged from the current research is that Incel culture is changing, and it is saturating more general and popularised youth cultures. Young people draw on materials, no longer from the alternative forums such a 4Chan (Nagle, 2017), but rather find in(cel)fluencers on popular platforms such as TikTok and YouTube. Content that one once needed to seek out Incel actively, as it was relegated to subversive platforms like 8chan or specific Incel forums on Discord, is now algorithmically offered, and often pushed, on popular teen platforms like TikTok. This is a new phenomenon, which we have termed Incel 2.0, in which Incel has become popularised and saturated into mainstream youth cultures, as was seen in much of the online fieldwork that was conducted during this study (exemplified in Figure 1).



Figure 1: Exemplar descriptions of analysed TikTok videos

Incel is, or rather was, by its very definition, about feeling involuntarily alone – a community of alone-ness (Ging et al., 2020). Previous engagements with this community have revealed stories of isolation, which at times escalated into frustration, anger, and discussions of revenge, but overwhelmingly, these stories were about loneliness. As one participant described 'they're missing that community. There is talk about rejection… Although it seems to be less romantic rejection. More based upon social rejection' (Regehr, 2020, p. 144). The fact that these cultures can now connect large numbers of young men is a distinctly new phenomenon. If *Incel 1.0* was centred around being an outsider. *Incel 2.0* is by contrast, often about camaraderie and utilising Incel and misogynistic rhetoric and tools to articulate collectively contemporary frustrations.

Some popular commentators have further attempted to situate the mainstreaming of what has previously been seen as a terrorist group (Hoffman et al., 2020), within the context of the post-Me-too era. That is, positioning the interest in Incel culture as a 'backlash' towards the Me-too movement (Equimundo, 2022), where conversations often bemoan cancel culture and the view that "woke-ism" has gone too far (Equimundo, 2022). This argument embeds the phenomenon within the broader growth of right-wing populism across Europe and the United States, and suggests that Incel culture is intrinsically linked to the alt-right (Ware, 2021) and popular forms of misogyny (Banet-Weiser, 2018; O'Neill, 2018), which in turn can lead extremism and to violent attacks (Banet-Weiser, 2018; Keskinen, 2013; Walton, 2012).

What is evident is that if *Incel 1.0* was about being a 'geek' (Massanari, 2017), or an outsider, *Incel 2.0* is about banding together, and finding empowerment through the misogynist language of Incel. As this report will further outline, through gaining traction on popular social media platforms, Incel terms and concepts have now found their way into schoolyard colloquialisms, and terms are being bandied about in school hallways. Here, we see a concerning pattern; how online behaviours and content are having tangible, negative impacts in offline interactions.

How online consumption impacts offline interactions

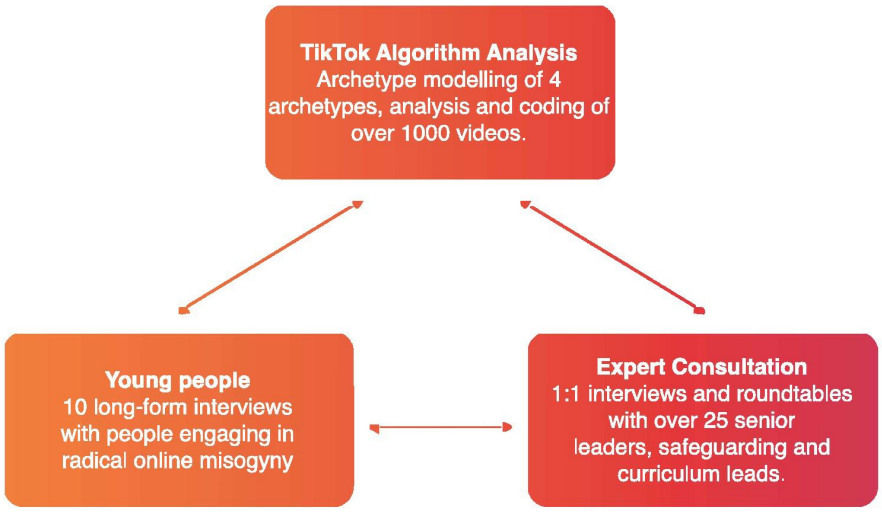
The impact of online consumption (particularly social media) for wellbeing and in young people's behaviours, has been one area that has been highlighted as particularly concerning, yet research has found it challenging to measure. Using large scale, social cross-sectional data from the millennium cohort study, Kelly et al., (2019), found that greater social media use was related to online harassment, poor sleep, low self-esteem, and poor body image, which in turn were related to higher likelihood of depression. High usage of social media has also been linked to social comparison and feedback seeking, which has been shown to be associated with higher depressive symptoms among adolescents (Nesi & Prinstein, 2015). Rounsefell et al., (2019) found that the impact of social media engagement on negative impact body image may have further direct impacts on food choices in some healthy young adults, particularly through exposure to 'fitspiration' posts and influencers on Instagram and TikTok. The further bleeding of online into offline behaviours can be seen in the increasing concerns around the role of access to and exposure to digital pornography, and its impact on violent sexual behaviours in young people's relationships (Children's Commissioner, 2023). As research by the Children's Commission found, 79% of young people had encountered violent pornography before the age of 18, and those who were frequent users of pornography were more likely to engage in sexual violence in their own relationships and interactions (Children's Commission, 2023).

However, online activity is complex and variable, varying from day to day, platform to platform and from person to person. This makes drawing direct links between social media and wellbeing more difficult. In an overarching review of six previous meta-analysis and systematic reviews, Odgers and Jensen (2019) found mixed results, with weak combinations of null, mixed, and positive associations regarding the impact of social media on mental well-being among young people. It is important to keep this in mind - for many young people, their online activity helps them find community, get a sense of belonging and interact creatively with their friends. However, like with many new and emergent technologies, critical approaches are needed in order to safeguard against those who are most vulnerable from being exploited.

To explore how themes of misogyny may be spreading online and manifesting in offline behaviours, we combine data from an algorithmic study with qualitative methods. Here, the algorithmic study was informed by qualitative interviews, with findings then cross referenced through interviews and roundtables with head teachers and safeguarding leads across the country. Our results highlight how educators believe there is a clear link between what young people consume on social media and the behaviour that they see playing out in classrooms and schools across the country.

METHODOLOGY

Our approach responds to gaps in our understanding of how the processes of social media platforms circulate and present material to individuals activity and interests, and of the impact of these on individual behaviours. This research draws on data triangulated across three sources, including (i) long-form interviews with young people, (ii) the analysis of over 1000 videos on social media and (iii) roundtables with school leaders to explore the perpetuation of themes of gender-based violence and misogyny among young people. Ethical approval was given for the study by the UCL Ethics Committee, and all participants gave informed consent to take part in the study.



(i) *Long form interviews with young people*

We undertook 10 long-form interviews with young people engaging with, and producing, online misogynist content. Working with a male researcher and documentary maker with previous experience of and contacts within within online Incel groups, participants were recruited from discord forums.

Discord and participant recruitment

Participant were recruited from the online platform Discord, which is a social media platform where users can communicate privately to each other or in virtual communities. There is no set makeup to discord forums. They are customisable and vary greatly and most rely on users moderating themselves. Some have quite strict rules, e.g. no pornographic material, no violent material, no racism. Others openly encourage distressing content. For example, in one server users were posting video clips of an animal being tortured. In another, there was a video of young man who had shot himself in front of a webcam. A lot of this content is real, not available on mainstream social media sites, and users enjoy posting it for shock value.

Discord allows you to create various channels within a forum. For example, a typical forum might have a channel called "rate me" where users post pictures of themselves and ask others to rate their looks out of 10. Some have quirkier channels, like place to post pictures of cats. Some have channels based around topics, like 'politics' or 'looksmaxing' or 'literature'. There are audio channels where users can join a group discussion, watch video content, or play video games together. Again, the topics of these group meetups varies greatly. We came across video channels where users were all watching an episode of My Little Pony. We also came across a video channel where a user was displaying and reading Mein Kampf.

To recruit young people engaging with and producing misognyst content, we focused on forums that were discussing content around women, relationships and Incel material. The researcher then joined these servers from one of their discord accounts. Many of these forums operated as simply gateways, where after joining, the researcher would be sent a link to a private server where users are openly discussing/sharing Incel content.

If accepted into channels, the researcher would message participants over a number of days, usually at night when they are online. Initial contact would also include a voice chat to describe the research and learn a bit more about their lives. These conversations also often focused around topics like loneliness and male mental health. Participants were most often recruited through a snowball sample approach on private forums. When a participant would respond and agree to an interview, they would often invite the researcher into a smaller server that they use with a closer group, where further approaches could be made.

Interviews were conducted either in an online audio call or over the phone, transcribed and then analysed thematically (Braun & Clarke, 2006; 2022). The results from interviews, alongside fieldnotes from online fieldwork working with this group, subsequently informed the development of the archetypes (outlined below).

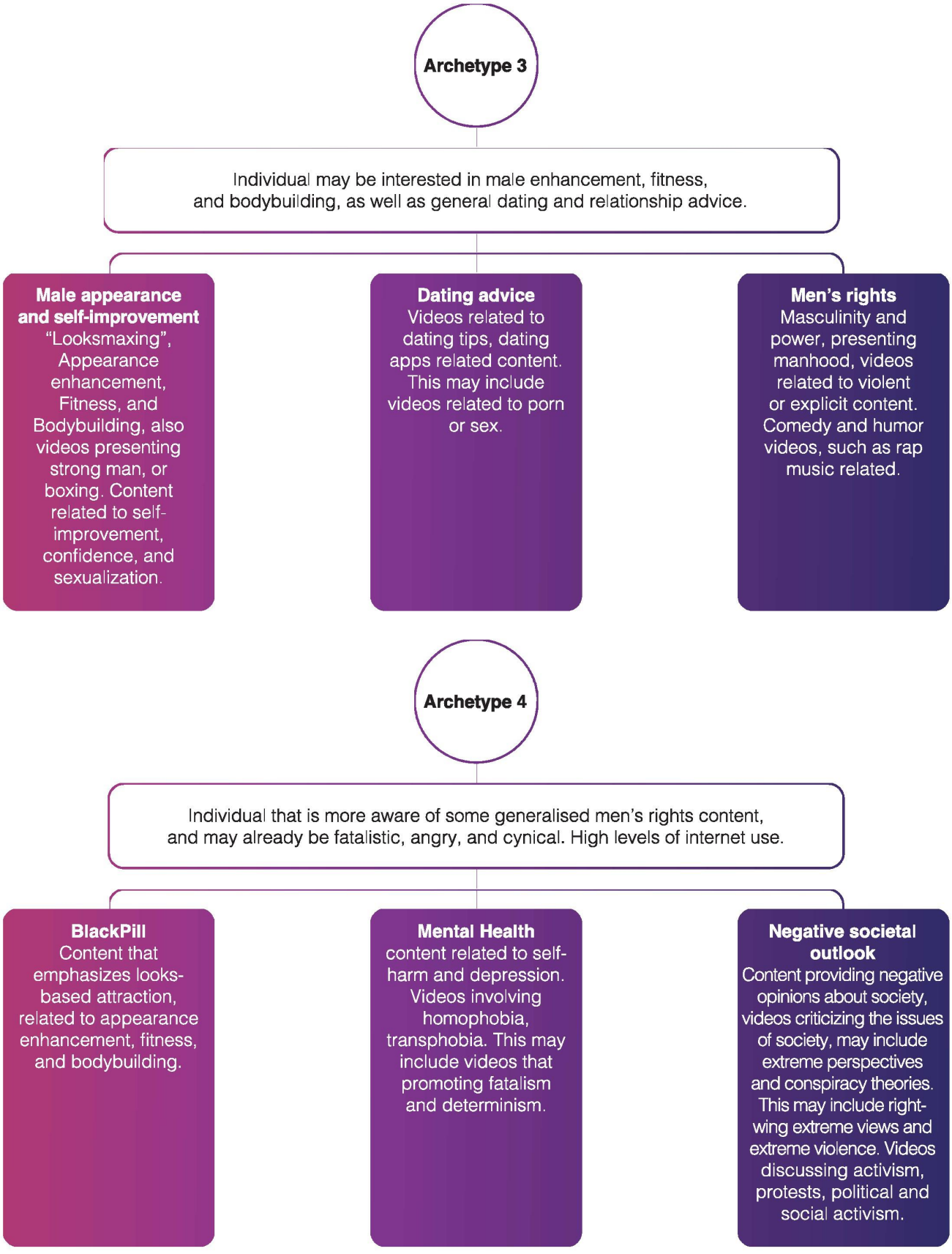
(ii) *TikTok Recommender Algorithms: Archetype Modelling*

To further understand the popularisation of misogyny, and the proliferation of Incel ideology onto online platforms, we undertook online fieldwork sought to explore how hypothetical individual users' journeys are exposed to materials online. Previously, algorithmic curation has posed challenges in accessing and analysing the online harmful content on social media platforms that are driven by recommender algorithms. One essential challenge is how to access such content on social media platforms which is driven by highly personalized recommendation algorithms. Unlike open forums like Reddit and 4chan, the harmful online content on TikTok, such as racial discrimination, extremely polarized political views, and sexism, is not easily visible as it has been deeply buried, and specifically allocated for some users who have shown potential interest in such content. The "For You" page on TikTok is highly personalised for each individual account, who is presented with videos and content by the people that account follows, or predominantly, recommended wider content from TikTok based on the individuals' inferred interests. How these interests are collected and inferred is driven by vast amounts of data collected by the algorithm, including (but may not limited to) the number of seconds a user stays on a particular video, whether they have liked or commented on similar content or hashtags used in their own generated content.

Additionally, there are restrictions for the public to access TikTok's video data, which also poses challenges for studying misogyny on TikTok. Although TikTok has announced Research API for academic researchers, scholars have argued that it has a large amount of limitations, such as only for researchers from non-profit universities in the U.S. and Europe, and limited reach of violative content (Kanthawala et al., 2022). Therefore, the current study introduced an approach named 'Archetype Modelling Methodology' to explore how algorithms influence the exposure to misogyny content to TikTok users. This follows similar procedures conducted by Amnesty International and Reset Australia to explore how TikTok's recommender system introduces content, including the radicalisation of that content, to users online.

Four archetypes were developed in our study to explore how initial interests in key theme areas influenced how, and to what extent, misogynistic content was presented to young people through the TikTok algorithm. These archetypes were built following from analysis our interviews with young people, and so were developed using specific terms and interests from the interviews, each with particular characteristics and vulnerabilities that were observed among those online communities.


Archetype 1
Individuals that are experiencing a sense of loneliness. They might have low self-esteem or low self-efficacy, experienced bullying and have high rates of internet-use. May be 'NEET' (not in employment, education or training).
Loneliness
This might include content that address feelings of loneliness, talk about bullying or school absence. This might include personal stories of resilience against bullying.
Anti-establishment
This might include content that criticize the capitalist system, capitalists, and politicians. This might include videos discussing socio-economic issues, inequality, and unemployment.
Male victimhood
Victimhood narratives and discussions regarding perceived societal biases against men. Content emphasizing the dominance of men, with rationalization, that is based on pseudoscience and statistics. Essentialistic explanations of gender hierarchy.
Masculinity and power
Inspirational stories of men, Career and Job Search Tips on how to make money.
Archetype 2
Individuals that are more focused on development of mental health knowledge and neurodiversity. They might seek out content that has an overtly negative focus and on negative neurodiverse experiences.
Neurodiversity
Content including a focus on personal stories and empathy about mental illness/ health, psychology, autism, ADHD or dyslexia.
Self-improvement narratives
Inspirational stories of people to achieve success, believing in yourself, content that has clear cut/black and white thinking and journeys out of adversity.
Negative societal outlook
Content providing negative opinions about society, videos criticizing the issues of society, may include extreme perspectives and conspiracy theories.


Archetype 3
Individual may be interested in male enhancement, fitness, and bodybuilding, as well as general dating and relationship advice.
Male appearance and self-improvement
"Looksmaxing", Appearance enhancement, Fitness, and Bodybuilding, also videos presenting strong man, or boxing. Content related to self-improvement, confidence, and sexualization.
Dating advice
Videos related to dating tips, dating apps related content. This may include videos related to porn or sex.
Men's rights
Masculinity and power, presenting manhood, videos related to violent or explicit content. Comedy and humor videos, such as rap music related.
Archetype 4
Individual that is more aware of some generalised men's rights content, and may already be fatalistic, angry, and cynical. High levels of internet use.
BlackPill
Content that emphasizes looks-based attraction, related to appearance enhancement, fitness, and bodybuilding.
Mental Health
content related to self-harm and depression. Videos involving homophobia, transphobia. This may include videos that promoting fatalism and determinism.
Negative societal outlook
Content providing negative opinions about society, videos criticizing the issues of society, may include extreme perspectives and conspiracy theories. This may include right-wing extreme views and extreme violence. Videos discussing activism, protests, political and social activism.

Following the identification of the four archetype's interests and the procedure flow for watching videos, (see Figure 1), each archetype was run consecutively, each for seven days between 22nd August and 27th September 2023 on four separate factory-reset iPads. No identification information was given to TikTok through account creation. Given the ethical limitations of interacting directly with users, particularly those under the age of 18, the only activity that was conducted on the archetype accounts was watching videos. No proactive activity, such as liking, commenting, or searching was conducted. For each archetype, a researcher watched videos on TikTok for one hour per day for seven continuous days. Based on the archetypes 'interested' list, if the videos presented through the algorithm were determined to align with the respective archetype's preference, they proceeded to watch the video to the end, and the shared link was documented. If the videos did not match the topics' preferences or were unrelated, they were skipped. All relevant videos from each day were documented and videos from day 2, day 5 and day 7 were subsequently thematically coded.

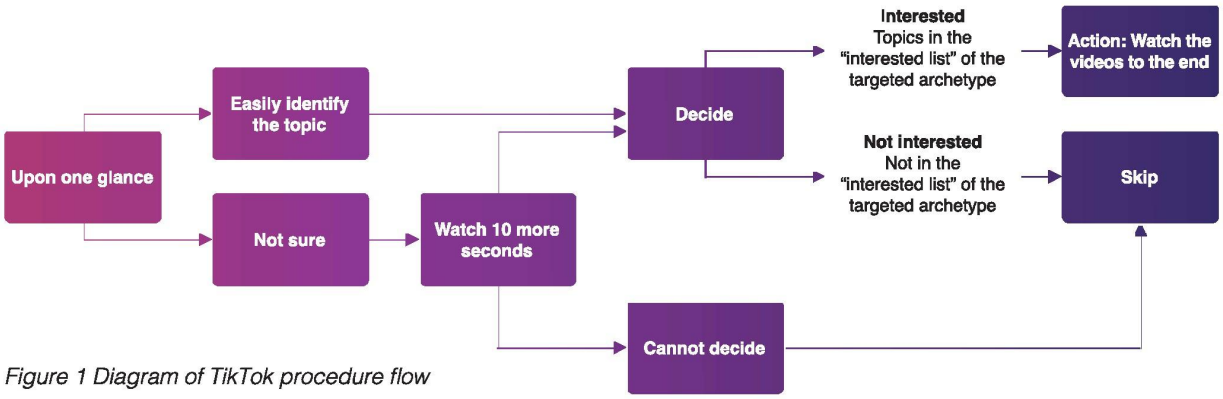


Figure 1 Diagram of TikTok procedure flow

Expert Consultation

The research team further undertook a series of nine interviews with safeguarding leads and members of senior leadership teams across the country to discuss the challenges facing schools and to explore how the trends that we had observed online were playing out in young people's relationships and behaviours. In collaboration with our partner, ASCL, the research team also conducted an expert roundtable event, attended by 25 school leaders, held in October 2023, to gather experiences on online misogyny and share best practice guidance around ways to tackle concerns across the curriculum. This qualitative data from the interviews, expert interviews and roundtables were analysed thematically, alongside the videos collected.

FINDINGS

1 Recommender algorithms and the gamification of radical ideology

Similar to the findings of Ribeiro et al. (2020), Reset Australia (2022) and Amnesty International (2023), our results demonstrated how the affordances of social media platforms, in this research evidenced by the recommender systems of TikTok, actively amplify and direct harmful content.

Our archetypes, informed by the experiences of young people interviewed in part one of the project, were built to reflect different types of interests that typical boys and men may express on these platforms. Throughout the algorithmic study, the research team tracked how each individual archetype interests and vulnerabilities were exploited by the algorithm; how over time they were presented with different radical content, and how the themes of that content changed as more time was spent watching and using the app. It was notable that young people themselves were acutely aware of these processes, and the impact that the changing content had over time:

> ‘It becomes dangerous when you’re on something [for] a specific niche. Because you’re getting everyone who’s in that same mind set digging each other up more and more and more and more until a point where like… you no longer your original yourself by the time you got to the end of it.’

We observed in particular how the algorithmic processes targeted different vulnerabilities of neurodiversity, loneliness, and mental health. Initial content often focused on material that recognised and sympathised with social difference, exploring themes of loneliness or self-improvement. However, as more time was spent on the app, this content increasingly focused on anger and blame. Alongside other mental health content, the content that the archetypes were exposed to transformed the ideologies of loneliness and difference into anger with an overtly misogynistic focus (Regehr, 2022). This echoes much of the research that has previously focused on Incel, where the use of ‘therapeutic’ language (Johanssen, 2023) allows them to strengthen group bonding, collectively vocalizing their shared issues of mental health and social anxiety, such as feelings of being ‘left out’ (Kay, 2021) in online platforms.

Through the algorithmic processes, we observed how these were played out in real time. After five days, **all archetypes saw a four-fold increase in the level of misogynistic content being presented on their “For You” page** (increasing from 13% misogynistic content to 56%), with those that were most focused on loneliness (Archetype 1) and radicalism (Archetype 4) seeing the biggest increase. The criteria of the for videos being misogynist referred to prior work (Shushkevich and Cardiff, 2019), which identified misogyny using the following criteria: ‘Stereotype & Objectification (a description of women's physical and/or comparisons to narrow standards), Dominance (an assertion of the superiority of men over women), Derailing (abuse of a woman), Sexual Harassment & Treats of Violence (actions as sexual advances, requests for sexual favors, harassment), Discredit (slurring over women with no other larger intention)’. If a video involved misogynistic content from any category, it was coded as such.

From day 5 to day 7, we saw a flattening out of this pattern, with misogynistic content decreasing slightly, whilst hyper masculinity content continued to rise (see Figure 2).

ASCL000007_0017

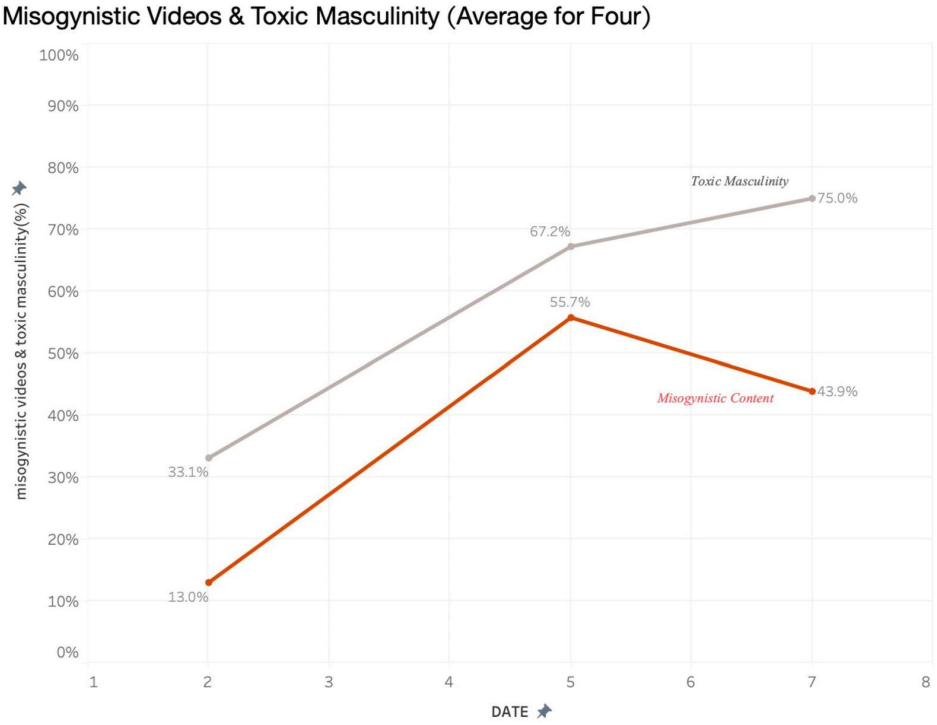


Figure 2 Percent of Misogynistic Videos and Toxic Masculinity content across all four archetypes

What was observed for two archetypes (Archetype 1 and Archetype 4) was that overtly and explicitly misogynistic content continued to rise across the seven days (Figure 3).

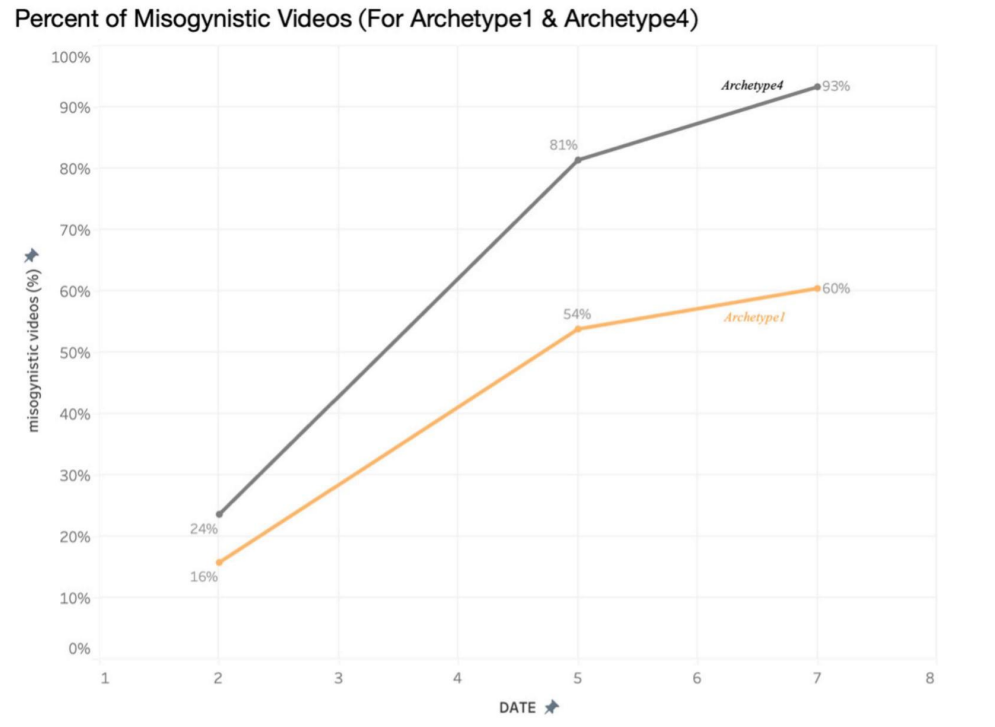


Figure 3 Percent of misogynistic content Archetype 1 & 4

For Archetypes 2 and 3, content that was framed around toxic assumptions of masculinity outstripped and began to replace more overly misogynistic content.

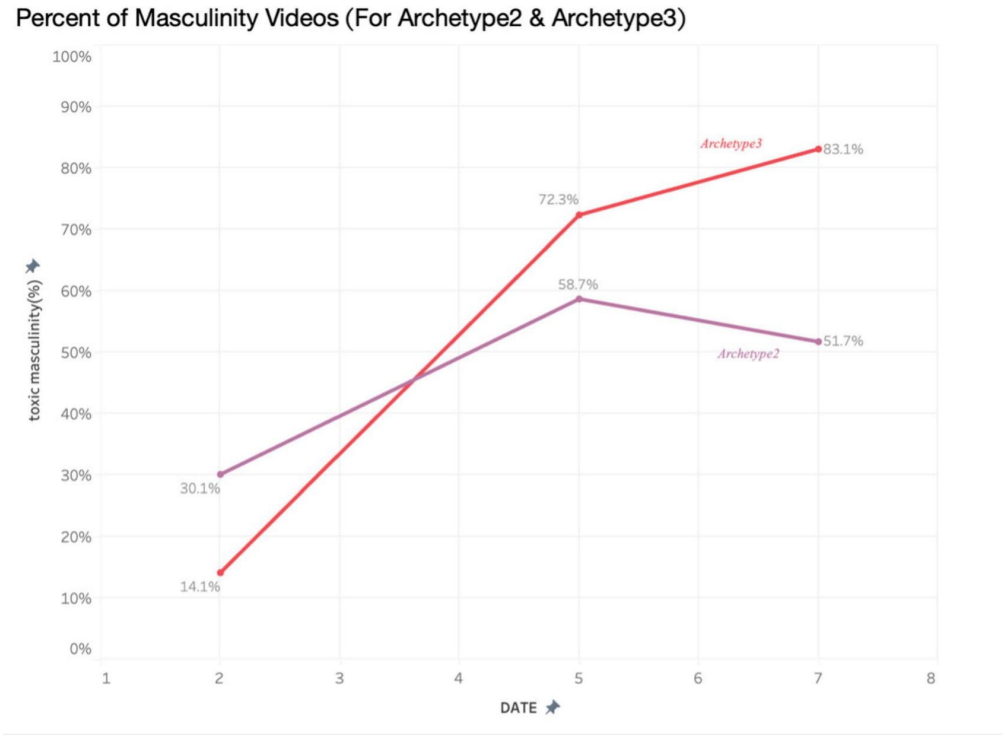


Figure 4 Percent of toxic masculinity content Archetype 2 & 3

Understanding patterns of indoctrination

To interrogate further how the algorithm transformed ideologies of loneliness or isolation into misogynistic or toxic masculinity content, we explored how different themes co-occurred and changed across the timeline of fieldwork. For Archetype 4, we observed how video content that centered around loneliness, life toughness and pressure on men in the first days was replaced by relationship advice, masculinity and how women think, all of which was filtered through an overtly misogynistic lens. This suggests that young people accessing content to deal with poor mental health or loneliness, are at high risk of further being pushed so called 'reasons' or targets of blame for their situation, which increasingly focus on women.

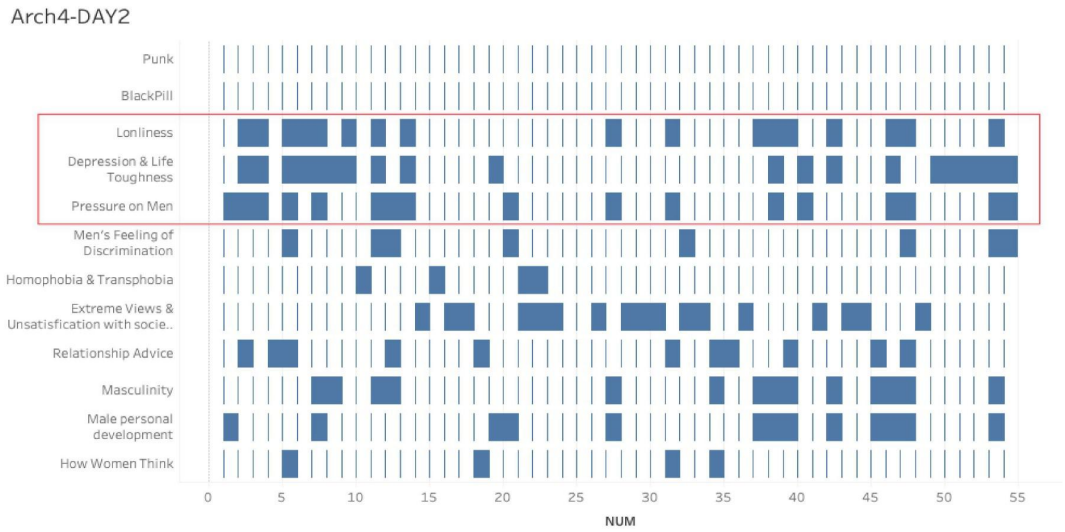


Figure 5 Thematic co-occurrences: Archetype 4, day 2

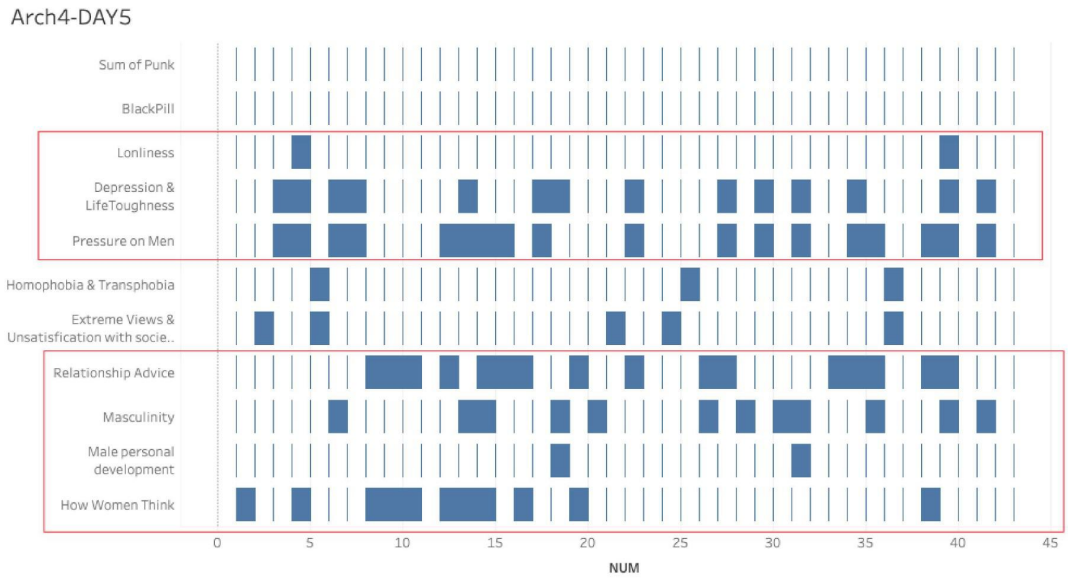


Figure 6 Thematic co-occurrences: Archetype 4, day 5

After day 5, we saw that the content that was once focused specifically on issues of loneliness or self-improvement has moved to focus more specifically on 'Pressure on Men' and 'Masculinity' but also co-occurred with content regarding relationship advice and how women think. This has the effect of drawing those two themes together, placing the responsibility and blame on women for the feelings of unfairness, disillusionment, and loneliness that young men in the content are expressing. This content, consumed at high dosages via the algorithm, has implications for young people's identity and belief systems.

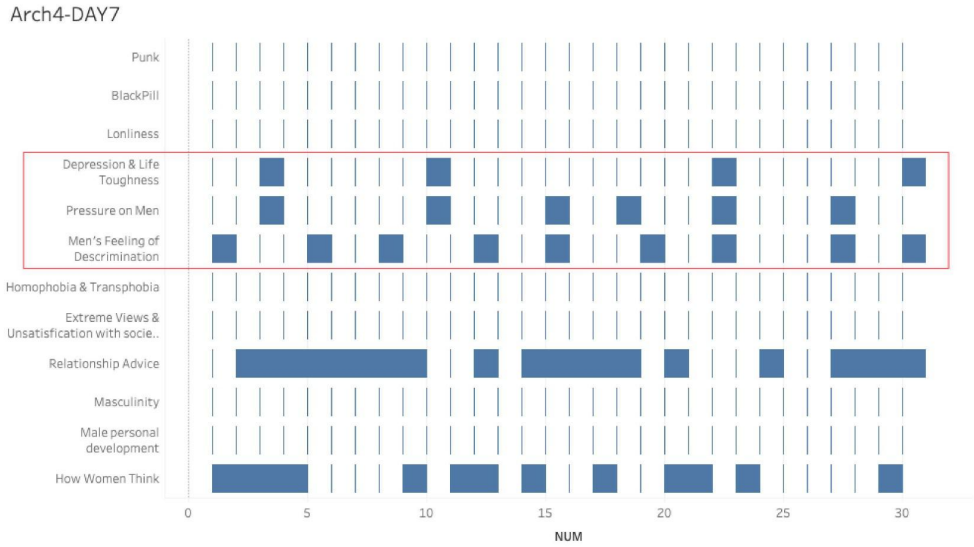


Figure 7 Thematic co-occurrences: Archetype 4, day 7

By day seven, as is evident by Figure 7, content of the videos now predominantly focusses on misogynistic content, particularly relationship advice and how women think. As one young person reflected:

> 'It's really getting targeted now and it's like […] obviously like cherry picked, […] but on TikTok like other women saying, okay, they'll only date a guy, if he's, if he's like this or like that.'

This young person is describing assumptions about female attraction to men based a narrow set of physical characteristics, which are 'like this or like that'. This sentiment directly echoes Incel ideologies regarding women's selection of male partners, which is often taken from eugenics concepts privileging whiteness and hyper masculine traits. This logic claims that women prioritise male physical appearance above all, and if an individual does not fit into these narrow physical standards, they can never improve their chances of having a meaningful relationship. The effect of consuming this concept for boys can be twofold. First that they develop prejudices against women for the supposed rejection, and further, that they feel they will be forever undesirable and alone.

Presentation of hyper-masculinity

Interestingly, for Archetype 2 and Archetype 3, we saw that overtly misogynistic content began to be replaced by content that focused on an extreme narrow presentation of hypermasculinity. For Archetype 3, this was particularly stark, where initial themes of relationship advice (see Figure 8), begin to be presented alongside masculinity and personal development content (see Figure 9), which is eventually overtaken by a focus on an extreme narrow presentation of hypermasculinity.

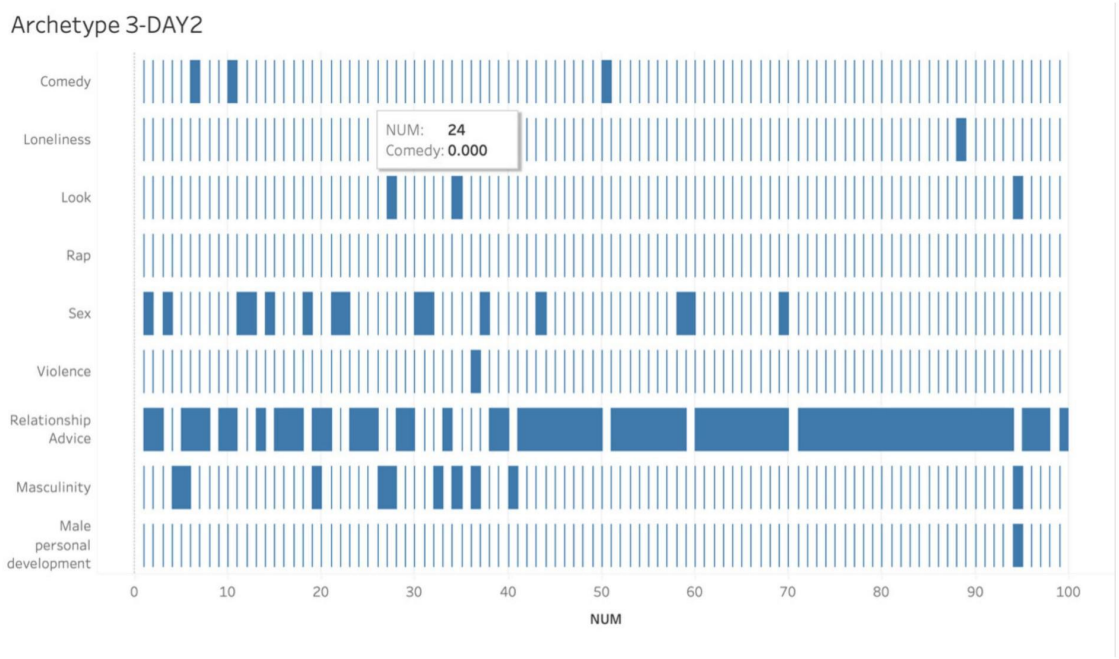


Figure 8 Thematic co-occurrences: Archetype 3, day 2

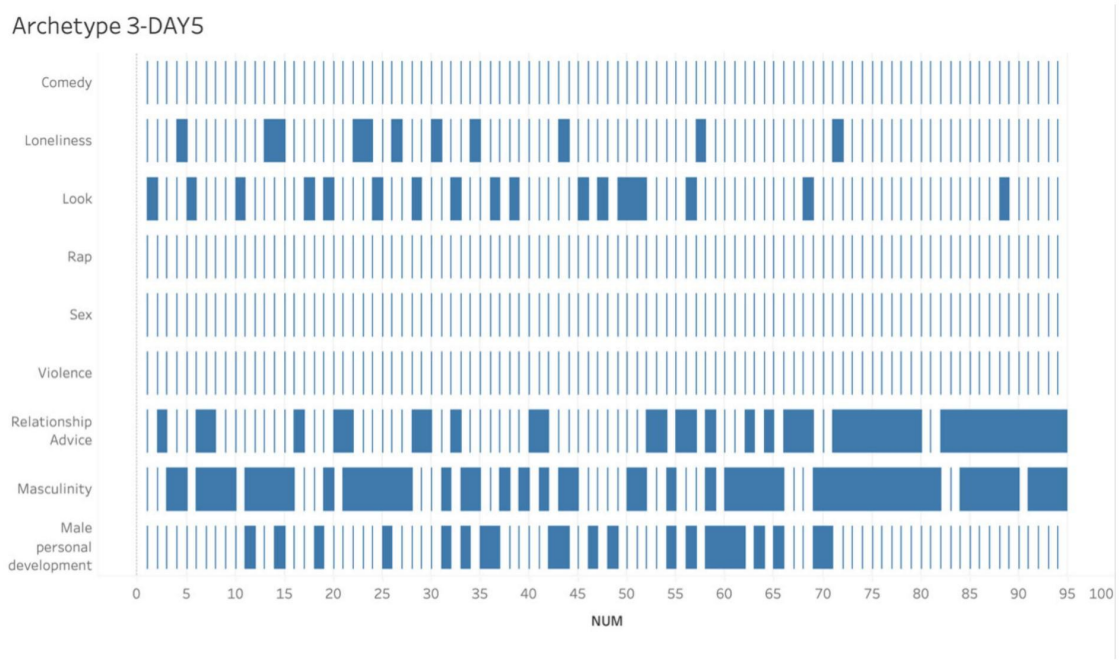


Figure 9 Thematic co-occurrences: Archetype 3, day 5

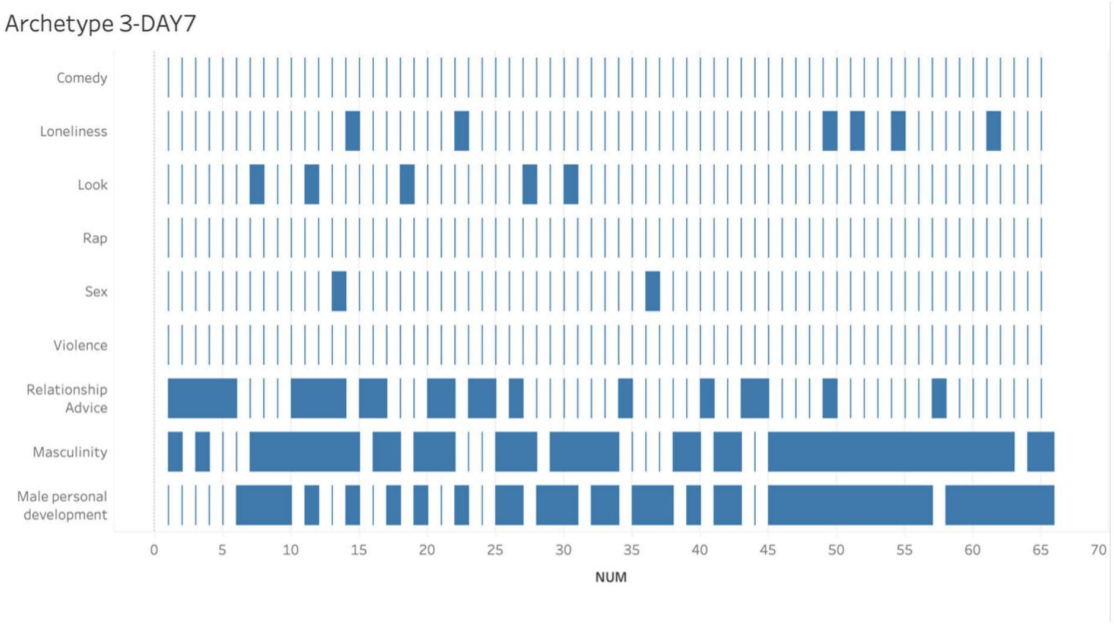


Figure 10 Thematic co-occurrences: Archetype 3, day 7

This new type of a hegemonic or hyper masculinity, was now directly linked to male sexual status, and emphasises how boys are pushed this narrow view of masculinity in order to answer the questions suggested by misogynistic content online, particularly those around loneliness and isolation.

In discussions with young people about their own journeys to radicalisation, individuals directly identified manosphere content as the initial area of interest that led to them exploring content further:

> '...you stump into manosphere comments, like there starts into red pill stuff and then (...) you'll probably get into black pill stuff (...)'

and were equally reflective of how their own exposure to radical content operated on a scale **'..there is a pipeline from Ben Shapiro to becoming a neo-Nazi...'**.

Here, these young people are clearly aware of the algorithmic process, or what they refer to as a 'pipeline'. To them, the pipeline is active in feeding them 'black pill stuff', which can consist of misogynistic, depressive and at times, self-harm, and mass violence material. This 'pipeline' is subsequently influencing them, and pushing them towards much more extreme versions of their original interests.

2 Incel 2.0 and the normalisation of radical, toxic misogyny

Overall, the analysis of the TikTok data and a review of the videos presented to the four accounts highlight the perpetuation of radical, toxic misogynistic content on TikTok. The sensitivity of the TikTok recommender systems meant that, once an account indicated a preference or watched a particular theme of videos, such as dating advice or 'how women think', their 'For You' feeds are dominated by this type of content. By the end of the seven days, much of the material presented to the four accounts had homogenised around similar themes of masculinity, self-improvement, dating advice and the expression of traditionalised roles of men and women.

Posts discussing dating and relationship advice

Across the reviewed content, the four accounts were dominated by posts that discussed dating and relationships, which included 'tips' on how to attract women, on the art of seduction, using stereotypical and pseudo-psychology to critique female behaviours, emphasising the importance of masculinity, strength, and self-improvement, as well as presenting traditionalised hierarchies and gender roles within relationships as aspirational. While some of this content was focused on male behaviour, much of this content progressed into specific videos that used negative, misogynistic tropes about 'how women think', with critical posts such as 'the truth about female nature…' and 'understanding the female narcissist' to present relationships as transactional, or with damaging stereotypes regarding expectations of sex.

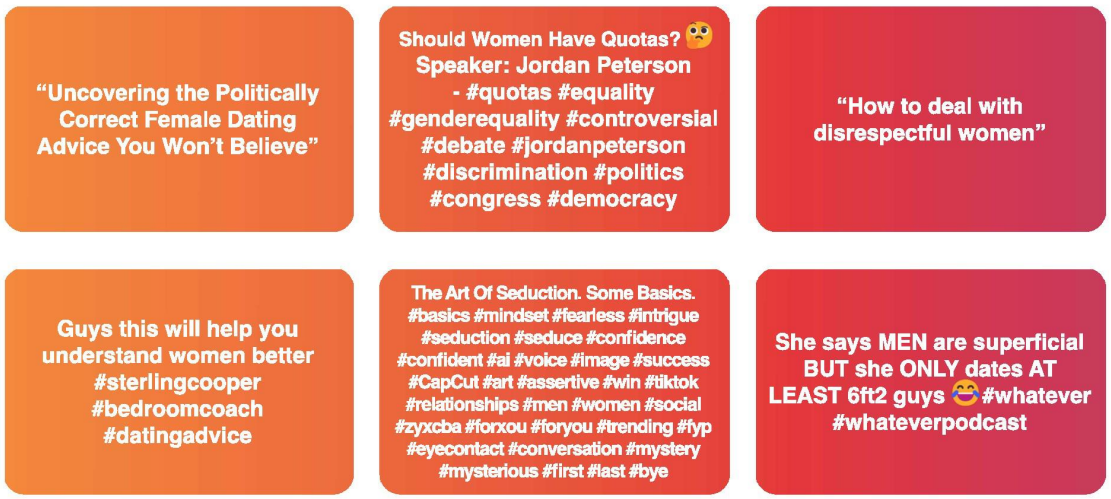


Figure 14 Exemplar descriptions of analysed TikTok posts featuring dating and relationship advice

These themes can be clearly linked to the Incel tropes of women having unrealistic physical standards, collectively desiring the most attractive alpha males, leaving other men behind in the sexual marketplace (Menzie, 2022, p.73). It is clear from this data that the prevalence of misogynistic content online has produced a new phenomenon of 'Incel 2.0' where young men are finding empowerment through the language of Incel and violent, toxic misogyny. Incel 2.0 utilizes misogynistic rhetoric and tools to articulate collectively contemporary frustrations.

Strength, masculinity and manosphere

Another theme across the video clips presented to the four accounts, often directly linked to dating advice, was the aspirational content that was associated with a highly narrow view of masculinity; one of physical strength, sexual prowess and extreme wealth.



Figure 15 Exemplar posts of Self-improvement and masculinity content

The relationship between the two themes of relationships and masculinity operated in a bi-directional feedback loop, with videos linked to masculinity and strength frequently co-occurring alongside those of relationship and dating advice. The intersection between these two themes was echoed by the accounts of young people, who were self-reflective and aware of these indoctrination processes:

> 'A lot of people you know, they're like, F**k, you know, I'm a teenager. Like, dude, I don't have a girlfriend, I'm trying to maximise…with like self-improvement...you sort of go down this rabbit hole, it happens… It's almost like political extremism…[and] you start off with like self-improvement and all that because you're insecure, you don't really get that many girls or you don't get girls at all…'

Here, this individual articulates how feelings of loneliness or fear of rejection are answered with a concept of maximising. 'Maxing' within the context of the manosphere refers to pushing social constructs like self-improvement, fitness, financial success to extreme levels. This individual even equates this pressure to 'maximise' with political extremism, where young people are indoctrinated this extreme self-help culture by way of the content they are consuming on social media.

Popularisation and behavioural impacts

Across the four archetypes, it was evident that the scale and framing of toxic misogyny online had resulted in the normalisation of misogynistic content and abuse. Analysis of the content emphasised how ideas that previously circulated on Incel platforms, notably those regarding male unfairness, 'looksmaxing' to achieve a physical ideal, and women's behaviours, are now widespread. The way that these viewpoints were presented, through 'soft' cultural forms such as inspirational content, memes, and parodies, served to mask the toxic and violent misogyny at its roots. As our interviews with young people further highlighted, this content is so pervasive that identifying points of indoctrination or influence is impossible; **'It's like memes, you can't really think 'when was the first time you saw this meme'. It's just everywhere...'**

School leaders further highlighted how they observed how this material was being played out in young people's behaviour. The normalisation of these kinds of ideas, through increasing exposure and micro-dosing online, led to a lack of recognition about their toxicity and impact.

> 'It's almost a lack of awareness, once you speak to the kids, they know it's wrong, but they almost don't realize it when they say it. And it's almost because they see it so much on social media, it's just normalized now'.

Others explicitly linked it to peer trends:

> 'I think for the vast majority of those boys, they don't actually believe what he's saying and fully follow it. But I feel like more, they're just following the normal doing it because they feel like that's what they do to fit in'.

As others reported, the growth in misogyny that had been observed in schools is a product of wider cultural problem regarding what is normal and increasing in technology and access to both misogyny online, and behaviours in the classroom. In relation to online sexual content specifically, one school leader noted:

> 'We did some work looking the number of students who have access to social media, and we're looking at the high 80s every year that have got access to this material online, so the link between the two (social media and misogyny) is definitely significant'.

This intersected with persisting challenges around digital sexual violence and image-based sexual harassment (IBSH), which schools are currently dealing with unsupported. As a school leader reflected when discussing with young people being sent unsolicited d**ck pictures online:

> 'It's happening out of school, so they're not reporting it. They don't, they are like 'Well, who cares? Like it's just one of those things'. So it's sort of battling a cultural shift, a social shift in just you know what's normal…So we've tried to work with them. No, no, that's not normal. That shouldn't happen. You shouldn't accept that'.

As they noted, existing reporting mechanisms may have had unintended effects, with young people explicitly underreporting it 'because it opens up this bigger, wider can of worms. So it's hard to know whether its [reduced], or they just think it's normal.'

What was most concerning across the reports from teachers, young people and through the video evidence collected is the pervasiveness of this content, and the way in which online content sanitised this material and presented it to young people in bitesize, digestible and entertaining chunks. This is having a micro-dose effect, saturating popular youth cultures with radical misogyny and toxic content which is having a hugely detrimental impact on the way young people interact with each other. What Incel 2.0 has quite clearly created, is a mass normalising of sexism and hate. Notably, this online engagement, just as in the case of Incel 1.0, is making young people angrier and unhappy. Currently, schools and parents are often left unsupported to deal with these issues.

3 Neurodiversity, Existing Vulnerabilities and Mental health

It was notable amongst the vast majority of the young people that we spoke to that mental health concerns seemed to co-occur with their exploration of more radical, misogynistic content online, echoing prior research that identified these online communities as a way of seeking companionship and a group identity (Preston et al., 2021). As one user described; 'I am lost (...) and directionless (...) I just want to fit in…my purpose is (…) mostly just to feel bad with other people who're feeling bad…'. This highlights how young people may not actively seek out these ideologies but rather, are being algorithmically offered hateful or misogynistic content, which in turn is being normalised by the high levels of their usage, increasing their dose.

Our findings in the algorithmic study highlighted a similar pattern, where content was initially focused on loneliness and a search for belonging and they were then pushed topics such hyper or hegemonic masculinity.

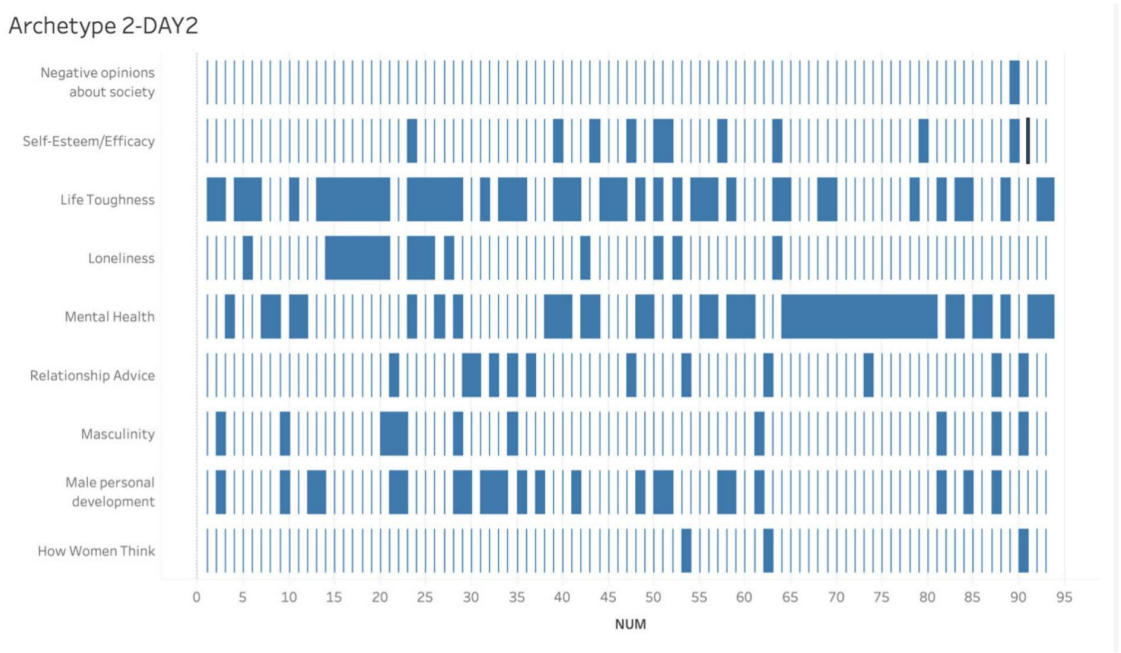


Figure 11 Thematic co-occurrences: Archetype 2, day 2

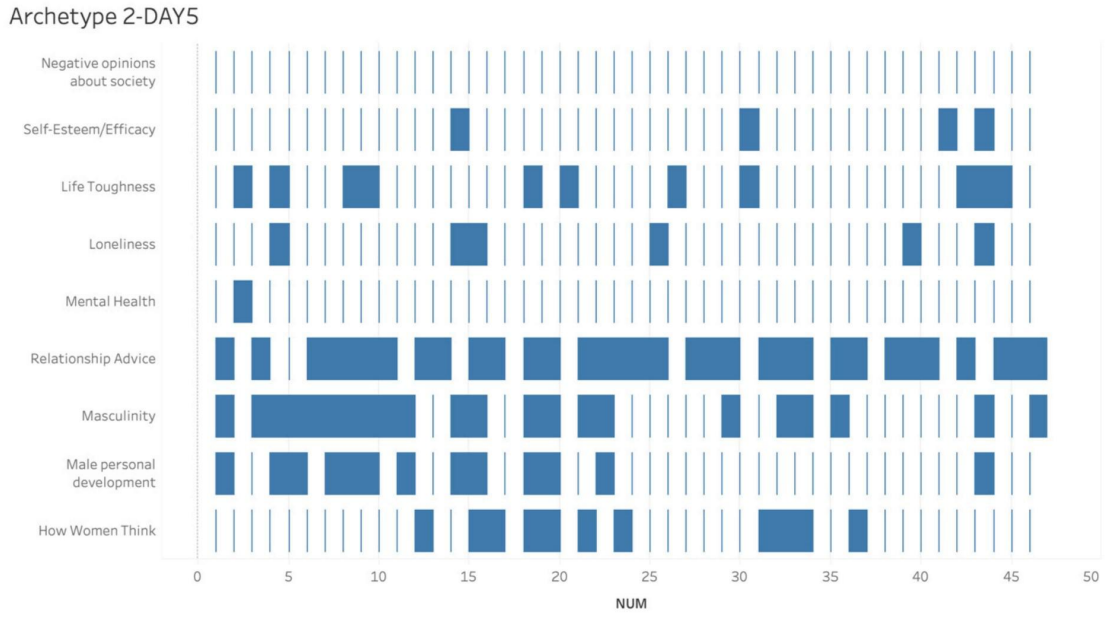


Figure 12 Thematic co-occurrences: Archetype 2, day 5

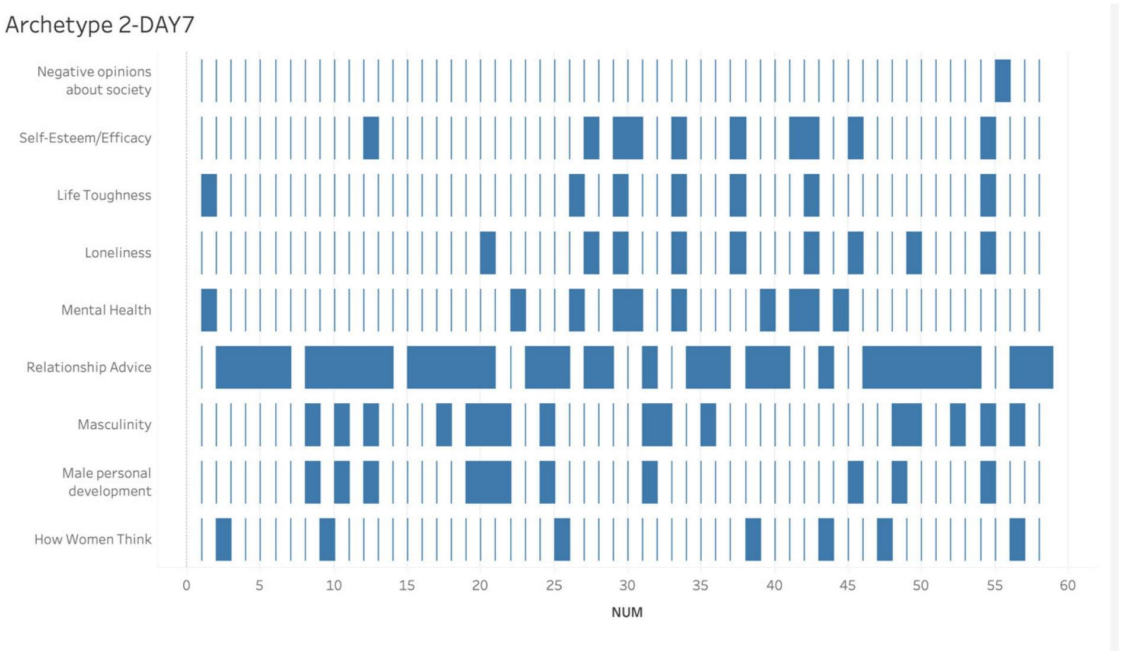


Figure 13 Thematic co-occurrences: Archetype 2, day 7

These patterns were echoed by the experiences of young people themselves who described their sense of isolation, the lack of awareness and the absence of supports; '…when I was in school there wasn't really enough done to address with these men, lonely men...'. Invalidation was also identified as a contributory factor to their situation (compounding their sense of hopelessness or anger). This suggests that a process is in place where young men, seeking community and understanding, or as an interviewee explains, a 'need for friendship and like for socializing' are finding an answer, endorsement and empathy in this toxic, harmful content.

As other studies have identified (Sparks et al., 2022; Costello et al., 2022), the most isolated and vulnerable individuals are those most at risk of indoctrination. This was echoed by one safeguarding lead:

> 'We've got a very small number of boys who are absolutely fixed on their views and we struggle to kind of get them to see anything else…And I think with regards to that, the more high risk students there, I think really that's where we do need some more specialist support within that. And I think that's something that I think for us at the minute in the city we're struggling with'.

This corresponds to concerns from teachers that neurodivergent adolescent men are a high-risk group and the reported prevalence of autism and neurodivergence in misogynist online communities (Tirkkonen & Vesperman, 2023). However, it is important to emphasise that this is a controversial topic with mixed results, calling for care and sensitivity to the complexities: 'we need to be 'extremely cautious when making generalizations and associating violence with autism' (Williams et al., 2021, pp. 395–396). There are a range of contested factors contributing to this association, which need to be considered in any endeavour to address and support. In the case of autism (as the most frequently cited example by community participants in our research and other studies), the experience of neuro-difference impacts on social relationships, cognitive processes (thinking and learning) emotions and sense of self, often affecting mental and physical health. The changes associated with adolescence and the transition to adulthood are particularly difficult for autistic young people, many of whom experience anxiety, depression and loneliness. Gender plays a critical role in identity formation and for neurodivergent cisgender boys seeking identification and confirmation of their masculinity, peer pressures to conform to social expectations of gendered behaviour are felt particularly strongly. Gender becomes a performative effort to fit in so much so that masking and assimilation are reported as having negative impacts on mental health for autistic adolescents (boys and girls), albeit with different behaviours associated with gender differences. While autistic girls are prone to internalising their differences (associated with underdiagnosis) (Mandy et al., 2012; Lai et al., 2015), boys are more likely to exhibit challenging or distressed behaviour, masking their mental health difficulties. All of this contributes to a toxic environment in which hate ideology can thrive (particularly misogyny) through difficulties with peers (e.g. bullying), female relationships, and body based anxieties (often compounded by sensory issues in autism).

Whilst neurotypical social contexts have been found to negatively shape neurodivergent formation of experience (Krueger, 2021a 2021b; Boldsen, 2022), online forums offer an alternative community and have been identified as positive for neurodivergent young people (Pavlopoulou, Ushe & Pearson, 2022) as a 'safe space' in which social interaction is easier and in which they can enjoy more agency and options for different forms of communication (e.g. use of chat functions, avatars, not having cameras on etc). As one participant explains, 'the internet…fulfils that need for friendship and like for socializing but doesn't fulfil that need for actual real-life interaction.'

These forums are often associated with focussed interests (e.g. gaming), enabling participants to experience a sense of belonging and sameness (rather than difference), all of which contribute to a sense of group identity, support emotion regulation and which fill an important need for young people who are struggling with neurotypical peer relationships. However, recent research (Tirkkonen & Vesperman, 2023) has identified the processes whereby online forums can have a more detrimental impact on neurodivergent young people, reinforcing a sense of shared hopelessness regarding life possibilities and contributing to their vulnerability to extremism and hate ideology. This creates a form of negative cognitive scaffolding in which the community echo chamber creates a form of emotional contagion, reinforcing and validating identity hostilities as experience so that extremes of thinking and feeling become increasingly entrenched as belief: 'the social interaction patterns in Incel communities provide a feeling of belonging at the expense of regaining hope' (Tirkkonen & Vesterman, 2023). The negative themes identified in the Tirkkonen and Vesperman study correspond to many of those identified in our interviews, particularly in terms of hopelessness, body image and futures. The complexities of these contextual and environmental factors are recognised by teachers:

> 'It's not necessarily all linked to Misogyny. I think it's more mental health issues. They're starting younger and we certainly saw a rise in kind of the Andrew Tate thing last year in boys…there was a pocket of boys that were talking about it all the time'.

4 Including young men in the conversation

In the review of the video content collected and school leaders responses to issues in schools, it became evident that existing approaches had limitations. The staff at one boys' school suggested they had been hit 'hard by Everyone's Invited,' the online campaign in which thousands of youths voiced experiences of gender violence. This staff member then questioned whether their now swift response to sexual assault allegations lacked nuance. They further wondered if their 'Incel problem' was a subsequent attempt on the boys' part, to swing the pendulum back. Instances like these highlight the problematic pedagogical and educational approaches available for working with boys that often 're-essentialise masculinities and embed limited assumptions about boys' (Equimundo, 2022, p. 38). Indeed, as one school leader reflected, previous pedagogical approaches had a negative effects regarding boys willingness to engage;

> 'I think whenever we talk about Andrew Tate, I think the defences come up and the boys think they're gonna get beaten with a stick again and we're trying to tell them how they should think, and how they should behave. But I think the more we flood them with the positive role models and raise aspirations around it, I think is a key strategy to use.'

Less prevalent are more nuanced discussions. That is, the ways in which Incel discourse might offer a means for young men to voice a fear of loss of control at a time that is very bleak for all young people. As one young person commented '...men are oppressed (…) isolated (…) I find some sort of solace in guys like Andrew Tate...'. Pedagogically embedded interventions, which address the ways in which we might include boys in the common goal of actualising healthy practices and relationships on and offline are more limited still. As one young man pointed out, 'I'm told everything I can't do and can't be' without being given positive role models or alternatives. It is in these gaps in understanding and educational approaches that we have entered the age of Incel 2.0 where extremist, misogynistic content has become saturated into popular youth ecosystems.

There is some admirable work being done in this area, such as, the new report The State of UK Boys: Understanding and Transforming Gender in the Lives of UK Boys (2022); and Education Scotland's, Mentors in Violence Prevention (MVP) program, which offers peer-to-peer learning where older boys mentor younger boys about gender violence and are currently developing Incel education. Nevertheless, important though these initiates are, they are only small turns of the dial. Tate is not going away. Though he is currently in prison and prior to his arrest, and had been banned from TikTok, his videos still circulate, inspire, and influence new forms of content creation on the platform. Further, there are a plethora of other influencers – or in(cel)encers - and content promoting misogynistic messaging to greater and lesser degrees that now populate the feeds of teenage boys. Thus, to focus on the specific influencer piece of digital content is to miss the point. Rather, this content is simply a symptom of a much larger cultural phenomenon: the popularising of technologically facilitated misogyny in the form of Incel 2.0 through mainstream social media platforms.

ASCL000007_0031

5 Promoting a healthy digital diet to support schools, parents, and young people

The problems highlighted in this report are emblematic of a wider gap in understanding healthy relationships to the digital space amongst youth and indeed, in the wider population. In combating this worrying trend, wider pressure is needed that continues to advocate for digital literacy training or how maintain healthy digital nutrition. As one school leader reflected, this requires an approach that includes parents and young people alongside the input that is already received in schools:

> 'We can teach them an hour, 2 hours, 3 hours a week. We can put it in assemblies, but that doesn't counteract the number of hours that they're exposed to things online. So then negative influences are far more than what we can do, you know. If I could just teach them about being safe online all day and I might have half a chance of tipping the balance a little bit on the positive things.'

As others recognised, a holistic approach where critical digital education was incorporated across the curriculum, rather than relegated to tutor time or group assemblies, was also needed:

> '[It's] the universal approach to it, I think is not something you can do once and done. It's gotta be something that's drip fed over time and certainly kind of discussed from many different aspects and the curriculum aspect of it is crucial as well'.

It is here that the inclusion of the voices of young people about their own experiences has been noted as particularly important. Young people create and co-create their online environments to address basic developmental needs, including identity formation, peer connection, and autonomy (Granic, Morita & Scholten, 2020), and are positive about the role of online communities to make them feel understood (RCPCH, 2019). But it is evident across these consultations with young people that many are acutely aware of the potential for harm that online platforms can present, and the risks for vulnerable individuals to be radicalised or exposed to further harmful content (such as 'Ana Coaches' (anorexia coaches), self-harm material and online misogyny) (RCPCH, 2019; Amnesty International, 2023). Indeed, as Vera San Juan et al. (2022) found in their consultations about future priorities for digital research, young people 'prioritized research questions related to social media and developing specific mental disorders, online bullying, and companies exploiting adolescents' vulnerabilities (for example through targeted publicity).' Within these discussions, evident again was less focus on the *time* that young people were spending online, but the *type* of content that they were exposed to and the affordances of online platforms which can direct them towards particular outputs and content creators.

School leaders were keenly aware that currently they have limited impact on what is happening online after school hours, and that parents don't always have the knowledge to support or engage with their child's online activity (or in some cases are perpetuating the same ideas and engaging with similar content). Therefore, alongside supports for teachers, awareness for parents needs to be improved, so they feel that they can engage with their children about their online behaviours.

Prohibitive approaches or a Healthy Digital Diet?

Prohibitive approaches to this problem, such as bans to social media or removing phones from schools, will have limited effectiveness. Recent evidence highlighting the ability of young people to circumvent age restrictions online (Farah & Milmo, 2023), and the rapid emergence of new or copycat platforms, such as BeReal, emphasises the limitations of legislation to respond in a digital age. Understanding how digital platforms now intersect with development and wellbeing requires a critical re-evaluation of how we understand digital engagement, with particular focus on the affordances of digital platforms in mediating and processing those interactions. As we have seen in the uses of social media across this report, there is a plethora of issues which are driven by the affordances of the platforms themselves. These relate to changes in personal relationships; interpersonal and romantic, u-loops and the lack of spontaneous cultural inputs which creates a more polarised society. We argue that the way in which this material is accessed - through ultra-processed platforms which alter and distort how content is consumed - leads to deeply unhealthy shifts in the way we think and interact with others.

Up until now, most of the guidance available focused on controlling the dose of digital engagement, or 'screen time'. However, 'screen time' doesn't account for the multitude of reasons young people use their screens, the advances in technology and the regularity in which phones are used. It is important to remember is that the core of social media usage is good. These elements include things like education, social interaction and creativity. However, the processing of certain material, by way of algorithms and content-creators, can in fact render it unhealthy or even harmful.

Moving away from prohibition, more fruitful, positive policy making should look towards the content of a health 'digital diet', which has been proposed to conceptualise the role of digital environments in development and public health for over a decade (Sieberg, 2011; The Wellbeing Thesis, 2023; Internet Matters, 2022; Orben, 2021). As Orben (2021) has argued, parallels between the study of food and technologies can support how we think about digital technology and its influence on development. 'For example, to understand diet we need to think about (a) what is being eaten, (b) the amount that is being eaten, (c) different food groups, (d) individual differences and (e) population differences', (Orben, 2021). However, much of the previous work on digital diets has, like in food and nutrition, been one of restriction (e.g. Sieberg's 2011 book "Digital Diet: The 4-Step Plan to Break Your Addiction and Regain Balance in Your Life). As this report has noted, digital platforms are now central and woven into young people's everyday lives, and the concept of limitation and restriction that has dominated nutrition discourse is now seeing a transition into looking more closely at the processes and ways in which substances that appear like food, are in fact 'ultra-processed' to the extent that they lose their nutritional value and instead fuel addiction and obesity.

Our argument is that similar to nutrition, the digital diet guidance needs to take a more nuanced approach. Like ultraprocessed foods, although at the core of social media usage are elements of educational, social, creative, and interactive material, the *processing* of that material, by way of algorithms and content-creators, can in fact render it unhealthy and potentially harmful. We are advocating for understandings of digital consumption to operate in the same way that we think about food consumption. To achieve this approach, investment into education is paramount, so that wellbeing of young people, and indeed the community at large is prioritised.

ASCL000007_0033

EVIDENCE BASED RECOMMENDATIONS

Advocates have repeatedly called to both hold social media companies accountable and prioritise educational initiatives, with Facebook whistle blower Frances Haugen stating that young people are often left to navigate social media driven issues on their own without adult support (Jenkin, 2022). Nevertheless, education-based interventions and support for teacher's concerns have been slow and woefully unable to keep pace with changing technologies and the implications they bring for young people. Blame is then often incorrectly placed on teachers to address issues, which primarily take place outside of school hours, or on the young people themselves. Our recommendations are both to hold social media platforms accountable, whilst also advocating for a healthy and critical digital education, which involves an iterative and student led approach to this problem. These proposed recommendations can only be taken forward with the support and investment of government. Without such investments, these issues will only become more entrenched and severe.

1 Holding Social Media Companies Accountable.

Previous research has examined how processes on social media amplify harmful content such as alt right (Riberio et al., 2020), manosphere content (Reset Australia, 2022), and self-harm material (Amnesty International, 2023). Our case study of online misogyny reveals the same processes at play, where hateful content is algorithmically offered to young people. In this research, we further trace how these online processes impact the school environment. In light of these evident behavioural and educational impacts, Big Tech companies need to be responsible for these harmful algorithmic processes. This means not just focusing on removing individual harmful content or videos, but on the underlying structures and processes that they have developed. Pressure needs to be applied so that big tech companies, like TikTok, address algorithmic harm and prioritise the wellbeing of young people over profit.

2 Implementing healthy digital diet education holistically across the curriculum.

We are advocating for an investment in critical digital literacy to be implemented holistically across the curriculum, ensuring the interventions are reactive to the needs of specific peer groups, up to date on new technological advancements, and further, places students as the co-creators of their safe practices for the online space in which they live.

This requires an ethical digital citizenship education and wider understanding of what makes healthy digital diet, which encourages young people to:

A. To think critically about the digital space and the content they consume within it.
B. To understand the impact of new technologies on their mental and physical health.
C. To engage all young people in the co-creation of new codes of conduct for the online world in which they live.

3 **Peer to Peer mentoring to support young people against gender violence and algorithmic harms.**

Inclusive and collaborative conversations are needed, as existing approaches exclude or isolate boys, and often amplify and entrench gender binaries. It is essential to include boys in discussions regarding online misogyny and developing critical digital literacy. This includes pedagogical interventions that champion youth voice, promote positive role models and are educative rather than providing a punitive response to boys' behaviours.

4 One approach, which helps integrate digital literacy across a school's social ecosystem, is peer-to-peer learning. This ensures that young people are included in these discussions and co-create new codes of conduct. An example, which has been developed alongside this research, is the Mentors in Violence Prevention (MVP) program pioneered by Angela McDonald and Lorna Aitken at Education Scotland. Here, older pupils are trained to lead sessions and mentor younger pupils around topics like online misogyny. This means that the older pupils are empowered into leadership roles, they are embedded in the school ecosystem and can continually monitor young students' progress[1]. This has the further impact of supporting a wider cultural change within schools and among young people.

Wider awareness of algorithmic process for parents and the community at large.

Finally, a recognition that parents also have role to play is needed. Schools are now met with extreme challenges around the impacts of social media consumption, but many times parents are not equipped with the knowledge, understanding or previous life experience to support young people. Often adults themselves are not aware of how algorithmic process function, or indeed their own social media addictions or biases, making parenting around these issues difficult. To effectively embed a healthy digital diet, we recommend a linked-up approach between public health, safeguarding, school leadership, teachers, and parents in order to support young people with key skills to recognize radicalisation and be critical about toxic online material as they transition to adulthood.

1 Evidence of the impact of MVP has been gathered through staff and pupil feedback, attitude questionnaires and focus groups, with both mentors and mentees reporting that they are more aware of the issues related to gender-based violence. In addition: (i) Staff report an increase in pupils who are ready to alert them to safety concerns. This increased communication allows staff to be proactive and to put support in place as necessary. (ii) Both pupils and staff refer to an improved ethos within school and a reduction of barriers between older and younger pupils. (iii) Mentors and staff identify a positive impact on mentors' confidence and the enhancement of a range of skills such as team-work and presentation skills. (MVP)

ASCL000007_0035

REFERENCES

American Academy of Child and Adolescent Psychiatry (AACP). (2020, February). *Screen Time and Children*. **https://www.aacap.org/AACAP/Families_and_Youth/Facts_for_Families/FFF-Guide/Children-And-Watching-TV-054.aspx**

Amnesty International. (2024). *Driven into Darkness.* **https://www.amnesty.org/en/documents/pol40/7350/2023/en/**

Boldsen, S. (2022). Autism and the sensory disruption of social experience. *Frontiers in Psychology*. 13 (874268) **https://doi.org/10.3389/fpsyg.2022.874268**

Children's Commission. (2023). *Evidence on pornography's influence on harmful sexual behaviour among children.* **https://assets.childrenscommissioner.gov.uk/wpuploads/2023/05/Evidence-on-pornographys-influence-on-harmful-sexual-behaviour-among-children.pdf**

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, *118*(9), e2023301118.

Costello, W. Rolon, V. Thomas, A.G. Schmitt, D. Levels of well-being among men who are Incel (involuntarily celibate). *Evolutionary Psychological Science*, 8 (4) (2022), pp. 375-390, **https://doi.org/10.1007/s40806-022-00336-x**

Daly, S. E., & Reed, S. M. (2022). "I Think Most of Society Hates Us": A Qualitative Thematic Analysis of Interviews with Incels. *Sex Roles*, *86*(1), 14-33.

Das, S. (2022, August 06). How TikTok bombards young men with misogynistic videos. *The Guardian*. **https://www.theguardian.com/technology/2022/aug/06/revealed-how-tiktok-bombards-young-men-with-misogynistic-videos-andrew-tate**

Dearden, L. (2023, February 17 2023). Counter-terror police concerned about violent misogyny and Andrew Tate rhetoric. *The Independent.* **https://www.independent.co.uk/news/uk/home-news/terrorism-incels-andrew-tate-police-b2283664.html**

Díaz, P. C., & Valji, N. (2019). Symbiosis of Misogyny and Violent Extremism. *Journal of International Affairs*, 72(2), 37-56.

Equimundo. (2022). *The State of UK Boys: Understanding and Transforming Gender in the Lives of UK Boys.* **https://static1.squarespace.com/static/5d77e56c1fc5e024160affa9/t/636a7734572b6e1fa5f279cd/1667921717569/State-of-UK-Boys-Long-Report.pdf**

Farah H,. & Milmo., D (2023, December 19). TikTok allowing under-13s to keep accounts, evidence suggests. *The Guardian.* **https://www.theguardian.com/technology/2023/dec/19/tiktok-allowing-under-13s-to-keep-accounts-evidence-suggests**

Floridi, L. (2016, November 29). Fake news and a 400-year-old problem: we need to resolve the 'post-truth' crisis. *The Guardian.* **https://www.theguardian.com/technology/2016/nov/29/fake-news-echo-chamber-ethics-infosphere-internet-digital**

Ging, D. (2019). Alphas, betas, and incels: Theorizing the masculinities of the manosphere. *Men and masculinities*, 22(4), 638-657.

Ging, D., Lynn, T., & Rosati, P. (2020). Neologising misogyny: Urban Dictionary's folksonomies of sexual abuse. *New Media & Society*, *22*(5), 838-856.

Granic, I., Morita, H., & Scholten, H. (2020). Beyond Screen Time: Identity Development in the Digital Age. *Psychological Inquiry*, *31*(3), 195–223. **https://doi.org/10.1080/1047840X.2020.1820214**

Hoffman, B. (2019). *How Serious Is White Nationalist Terrorism?* **https://www.cfr.org/in-brief/how-serious-white-nationalist-terrorism**

Hoffman, B., Ware, J., & Shapiro, E. (2020). Assessing the threat of incel violence. *Studies in Conflict & Terrorism*, 43(7), 565-587.

Horne, M. (2021, August 20, 2021). Teachers in Scotland will be trained to spot incel extremists. *The Times*. **https://www.thetimes.co.uk/article/teachers-in-scotland-will-be-trained-to-spot-incel-extremists-nvl8dz572**

Internet Matters. (2022). **https://www.internetmatters.org/wp- content/uploads/2023/01/Internet-Matters-Balancing-Screen-Time-guide-Jan23.pdf**

Jenkin, E. (Director). (2022). *The Instagram Effect* [Documentary Film]. British Broadcasting Corporation.

Johanssen, J. (2023). Reconsidering trauma and symbolic wounds in times of online misogyny and platforms. *Media, Culture & Society*, *45*(1), 191- 201. **https://doi.org/10.1177/01634437221127362**

Kanthawala, S., Cotter, K., Foyle, K., & DeCook, J. R. (2022). It's the Methodology For Me: A Systematic Review of Early Approaches to Studying TikTok. In T. X. Bui (Ed.), *Proceedings of the 55th Annual Hawaii International Conference on System Sciences, HICSS 2022* (pp. 3105-3121). IEEE Computer Society.

Kay, J.B. (2021). Abject desires in the age of anger Incels, femcels and the gender politics of unfuckability. Mary Harrod, Suzanne Leonard, Diane Negra (*Eds*) in *Imagining "We" in the Age of "I" Romance and Social Bonding in Contemporary Culture*. Routledge, London.

Kelly, Y., Zilanawala, A., Booker, C., & Sacker, A. (2018). Social Media Use and Adolescent Mental Health: Findings From the UK Millennium Cohort Study. *eClinicalMedicine*, *6*, 59–68. **https://doi.org/10.1016/j.eclinm.2018.12.005**

Krasodomski-Jones, A. (2017). *Talking To Ourselves? Political Debate Online and the Echo Chamber Effect*. Demos. **https://demos.co.uk/wp-content/uploads/2017/02/Echo-Chambers-final-version.pdf**

Krueger, J. (2021a). "Agency and atmospheres of inclusion and exclusion" in Trigg, D (Ed). *Atmospheres and Shared Emotions*. (New York: Routledge), 111–132.

Krueger, J. (2021b). Finding (and losing) one's way: autism, social impairments, and the politics of space. *Phenomenology of Mind. 21*, 20–33. **https://doi.org/10.17454/pam-2102**

Lai, M-C., Baron-Cohen, S., & Buxbaum, JD. (2015). Understanding autism in the light of sex/gender. *Molecular Autism. 6*(1) **https://doi.org/10.1186/s13229-015-0021-4**

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018, March 09). The science of fake news. *Science*. **https://www.science.org/doi/10.1126/science.aao2998**

Lewis-Kraus, G. (2022, June 03). How Harmful is Social Media? *The New Yorker*. **https://www.newyorker.com/culture/annals-of-inquiry/we-know-less-about-social-media-than-we-think**

Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012) Sex differences in autism spectrum disorder: Evidence from a large sample of children and adolescents. *Journal of Autism and Developmental Disorders*. *42*(7), 1304–1313. **https://doi.org/10.1007/s10803-011-1356-0**

Massanari, A. (2017). #Gamergate and The Fappening: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society*, *19*(3), 329-346.

Matamoros-Fernández, A., Gray, J. E., Bartolo, L., Burgess, J., & Suzor, N. (2021). What's "Up Next"? Investigating Algorithmic Recommendations on YouTube Across Issues and Over Time. *Media and Communication*, *9*(4), 234–249. **https://doi.org/10.17645/mac.v9i4.4184**

Menzie, L. (2022). Stacys, Beckys, and Chads: the construction of femininity and hegemonic masculinity within incel rhetoric. Psychology & Sexuality, 13:1, 69-85. **https://doi.org/10.1080/19419899.2020.1806915**

Nagle, A. (2017). *Kill all normies: Online culture wars from 4chan and Tumblr to Trump and the alt-right*. John Hunt Publishing.

Nesi, J., & Prinstein, M.J. (2015). Using Social Media for Social Comparison and Feedback-Seeking: Gender and Popularity Moderate Associations with Depressive Symptoms. *Journal of Abnormal Child Psychology, 43*, 1427–1438. **https://doi.org/10.1007/s10802-015-0020-0**

Odgers, C. L., & Jensen, M. R. (2020). Annual Research Review: Adolescent mental health in the digital age: facts, fears, and future directions. *Journal of Child Psychology and Psychiatry*, *61*(3), 336–348. **https://doi.org/10.1111/jcpp.13190**

Ofcom. (2021). *Children and parents media use and attitudes report 2020/21*. **https://www.ofcom.org.uk/__data/assets/pdf_file/0025/217825/children-and-parents-media-use-and-attitudes-report-2020-21.pdf**

Ofsted. (2021). *Review of sexual abuse in schools and colleges*. **https://www.gov.uk/government/publications/review-of-sexual-abuse-in-schools-and-colleges/review-of-sexual-abuse-in-schools-and-colleges**

Orben, A. (2022). Digital diet: A 21st century approach to understanding digital technologies and development. *Infant and Child Development*, *31*(1), e2228. **https://doi.org/10.1002/icd.2228**

Papadamou, K., Zannettou, S., Blackburn, J., De Cristofaro, E., Stringhini, G., & Sirivianos, M. (2021). " How over is it?" Understanding the Incel Community on YouTube. *Proceedings of the ACM on Human-Computer Interaction,* 5(CSCW2), 1-25.

Pavlopoulou, G., Usher, C., & Pearson, A. (2022). 'I can actually do it without any help or someone watching over me all the time and giving me constant instruction': Autistic adolescent boys' perspectives on engagement in online video gaming. *British Journal of Developmental Psychology. 40*(4), 557-571. **https://doi.org/10.1111/bjdp.12424**

Preston, K., Halpin, M., & Maguire, F. (2021). The black pill: New technology and the male supremacy of involuntarily celibate men. *Men and Masculinities, 24*(5), 823–841. **https://doi.org/10.1177/1097184X211017954**

Regehr C., Glancy G., Ramshaw L., & Carter A. (2017). A comprehensive approach to managing threats of violence on a university or college campus. *International Journal of Law and Psychiatry, 54*, 140–147.

Regehr, K. (2020). In (cel) doctrination: How technologically facilitated misogyny moves violence off screens and on to streets. *New Media & Society*, 1461444820959019.

Reset Australia. (2022). *Algorithms as a weapon against women: How YouTube lures boys and young men into the 'manosphere'.* **https://au.reset.tech/uploads/algorithms-as-a-weapon-against-women-reset-australia.pdf**

Ribeiro, M. H., Blackburn, J., Bradlyn, B., De Cristofaro, E., Stringhini, G., Long, S., ... & Zannettou, S. (2021, May). The evolution of the manosphere across the web. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 15, pp. 196-207).

Ringrose, J., Regehr, K., & Milne., B (2021) *Understanding and Combatting Youth Experiences of Image-Based Sexual Harassment and Abuse.* **https://www.ascl.org.uk/ASCL/media/ASCL/Our%20view/Campaigns/Understanding-and-combatting-youth-experiences-of-image-based-sexual-harassment-and-abuse-full-report.pdf**

Rounsefell, K., Gibson, S., McLean, S., Blair, M., Molenaar, A., Brennan, L., Truby, H., & McCaffrey, T. A. (2020). Social media, body image and food choices in healthy young adults: A mixed methods systematic review. *Nutrition & Dietetics*, *77*(1), 19–40. **https://doi.org/10.1111/1747-0080.12581**

Royal College of Paediatrics and Child Health (RCPCH). (2019). *The health impacts of screen time: a guide for clinicians and parents.* **https://www.rcpch.ac.uk/sites/default/files/2018-12/rcpch_screen_time_guide_-_final.pdf**

Sieberg, D., (2011) *Digital Diet: The 4-Step Plan to Break Your Addiction and Regain Balance in Your Life: The 4-step plan to break your tech addiction and regain balance in your life.* Three Rivers Press.

Shushkevich, E., & Cardiff, J. (2019). Automatic misogyny detection in social media: A survey. *Computación y Sistemas*, 23(4), 1159-1164.

Sparks, B., Zidenberg, A. M., & Olver, M. E. (2023). One is the loneliest number: Involuntary celibacy (incel), mental health, and loneliness. *Current psychology*, 1–15. Advance online publication. **https://doi.org/10.1007/s12144-023-04275-z**

Speckhard, A., & Ellenberg, M. (2022). Self-reported psychiatric disorder and perceived psychological symptom rates among involuntary celibates (incels) and their perceptions of mental health treatment. *Behavioral Sciences of Terrorism and Political Aggression*, 1-18.

UCL

University of Kent

JANUARY 2024