University of BRISTOL

DEPARTMENT OF COMPUTER SCIENCE

Timeline Curation

Safety and Wellbeing Practices of Twitter Users

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Bachelor of Science in the Faculty of Engineering.

Thursday 4th May, 2023

Abstract

Social media is a valuable asset to everyday life and provides great enjoyment to many people - allowing connectivity across the globe and bringing people closer together regardless of physical distance. However, it is rife with anti-social behaviour and intentionally unpleasant content, ranging from spam to targeted harassment.

On websites such as Twitter and Instagram, which make use of a "timeline" feature, users have the ability to avoid this content to some degree by using built-in timeline curation tools like blocking users and hiding posts, but the impact and usefulness of these tools for users is not currently well explored.

Through semi-structured interviews with 12 participants, this study collects data on the day-to-day experience of Twitter users, their experiences with disagreeable content and the timeline curation tools they make use of. Qualitative analysis of our results reveals that users tend towards additive rather than subtractive curation, preferring to be selective about what content they add to their feed rather than take steps to remove it. The analysis identifies themes of users' distrust of the algorithm and concerns about overly-curated experiences leading to an "echo chamber" effect. Our findings will provide more insight into how users interact with timeline curation tools, so that they can be improved with the users' concerns in mind.

Contents

List of Figures

List of Tables

Ethics Statement

An ethics application for this project was reviewed and approved by the faculty research ethics committee as application 13425.

Supporting Technologies

- I made use of Microsoft Teams' built-in transcription tools to perform a "first-run" transcription of my interviews
- I used NVivo, a Computer-Assisted Qualitative Data Analysis Software, to support qualitative analysis and coding of my data

Chapter 1

Introduction

Social media websites have developed considerably since their initial rise to popularity, with people making use of them frequently in their day-to-day life. However, despite the initial intentions of their creators, many sites are prone to antisocial behaviour, harassment, and other risks to users' enjoyment and wellbeing ([53],[87], [24]). Twitter has been identified as a platform where this kind of behaviour is particularly prevalent. In 2015, its own chief executive acknowledged that the site "sucks at dealing with abuse and trolls" [35].

In order to allow its users better control over what they see on the site and improve user experience, Twitter offers a number of built-in timeline management tools. These include the ability to follow [80], block [82], and mute [83] other accounts and topics, as well as the functionality to report inappropriate content to the site's moderation team [84].

Little investigation has been done into how successful these tools are in reducing the amount of unpleasant content users are exposed to, or even how much they impact the user's day-to-day experience.

Through semi-structured interviews with twelve participants, this research paper investigates how these participants make use of timeline curation tools on a day-to-day basis, as well as how their experience as a whole is affected by these tools. Qualitative analysis reveals that users' Twitter experience is greatly shaped by their choices of timeline settings, and identifies trends of users preferring to additively, rather than subtractively, shape the content they see.

Through understanding current user behaviour and opinions on the tools they use, this paper offers insights for future changes to Twitter's timeline management tools and a user-informed basis on which they can be improved.

1.0.1 Aims and Objectives

1. Review literature about online safety, harm, wellbeing on Twitter and other social media

2. Conduct qualitative interviews with regular Twitter users

3. Transcribe interviews and analyse data using thematic analysis

1.0.2 Research Questions

1. How do users make use of Twitter's existing timeline curation tools, and do these tools fit the users' needs?

2. How does Twitter's algorithmic timeline affect the quality of experience for users?

Chapter 2

Background

2.1 Defining Digital Wellbeing

In order to gauge the effect of timeline management tools and the ability to curate their social media experience on the user's wellbeing, we must first consider how "wellbeing" is defined in a digital sense. Digital wellbeing is a concept that has been discussed by a number of studies. Mariek [88] defines it as "a subjective individual experience of optimal balance between the benefits and drawbacks obtained from mobile connectivity," while Burr et al [15] describe it as "the impact of digital technologies on what it means to live a life that is good for a human being in an information society", and Bartsch defines it as "an umbrella term for a state of individual wellbeing that is supported by digital technologies." [4]

In this paper, we operate from a definition of digital wellbeing as a concept that incorporates both the individual and societal experience, maximising positive impact and individual agency while reducing the drawbacks of engaging with digital technologies.

In the sense of defining digital wellbeing in regards to timeline curation, the key point is that a user should be able to control what they see on a given timeline-based site in order to see the content they are interested in and reduce their exposure to content they don't want to see – "disagreeable content". For the purposes of analysing users' responses to different kinds of content they may be exposed to, I have identified three categories of "disagreeable content."

1. Trolling

   "the posting of messages via a communications network that are intended to be provocative, offensive or menacing" [7]

2. Harassment

   "behaviour towards someone that is threatening or that annoys or upsets them" [63]

   I distinguish harassment from trolling by considering whether or not it is targeted – users might react differently to an offensive comment directed at them than to one with a less specific message.

3. Misinformation

   "false information that is spread, regardless of whether there is intent to mislead" [25]

While these types of content likely have different causes and provoke different responses in a social media user, they can all be considered as "content a user would wish to avoid", and so I categorise them as three categories of the same issue.

Trolling has undergone a change in meaning over time, shifting from a more "mischievous", but ultimately harmless, practice to one with more malicious intent [8]. Hannan [32] describes how "trolling" has gone from obscurity to a more mainstream practice, often no longer hidden behind anonymity on social media. He identifies how logical and rational arguments are now frequently given less attention than "short, biting sarcasm".

Harassment is highly prevalent – forty percent of adult internet users in the US have personally experienced some form of online harassment [27] – and can have highly harmful impacts on users of social media. Pater et al [57] investigate how different social media platforms form policies against harassment, considering not just how these platforms define harassment but what actions they take against it. Jhaver et al [36] take a different approach, gaining a perspective on online harassment by observing behaviour on and interviewing members of an internet community notorious for cyber-harassment.

Misinformation is a potential risk to users on most social media sites. Tran et al [77] investigate the role social media plays in misinformation during humanitarian crises, but with the wide range of purposes social media is used for, it also plays a part in misinformation on political matters, such as the 2016 US presidential election [2], and has been observed to be highly prevalent in the area of health misinformation [89]. There have been suggested means of reducing the impact of misinformation, such as using Facebook's related links feature to counter shared articles containing misinformation. [12]

As in a timeline-based social media, content is displayed to a user through the accounts and topics they "follow", the ability a user has to make active choices regarding which things they do follow affects their digital wellbeing. However, social media sites increasingly make use of an "algorithmic" timeline, where content is selected and prioritized based on factors decided by the site. On Twitter, the algorithm is affected by factors such as the popularity of a post or its conformation to certain traits associated with a post being favourably received by the viewer [81]. With this in mind, the algorithmic feed and how it operates also affects a user's experience and should be considered. There has been comparison [3] of the content algorithmic and chronological timelines show to users, but this study does not focus on the personal experience of users and identifies a need for a sample of real-world users rather than archetypal agents. Saurwein et al [67] have performed a broad investigation of the link between algorithms and harms on social media platforms, confirming that algorithms on sites such as Facebook, YouTube and Twitter have a considerable role in digital harms.

2.2 Digital Wellbeing and Twitter

Twitter in particular has been identified as a platform where disagreeable content and anti-social behaviour is particularly prevalent [35]. There are a number of reasons hypothesised for why this is the case: because "bad news" attracts more engagement [33], because Twitter's reduced character limit does not allow for nuance [55],or because Twitter encourages users to speak impulsively. [55] This behaviour seems to be on the rise, with the proportion of users suspended by Twitter increasing dramatically. [39]

A large proportion of sites of this kind make use of "timeline" functionality, where the user scrolls through a feed of posts from other users – some sites allowing a variety of content, and others narrowing down the medium. Instagram, for example, focuses on the sharing of images and videos, while Twitter was initially intended to feature only short text posts of no more than 140 characters though it has now broadened its used post types to include photographs, videos and longer-form Tweets.

Masciantonio et al [41] investigate Twitter and a number of other social number sites that make use of timelines in their effects on wellbeing. Active and passive use of each site (Facebook, Instagram, Twitter and TikTok) were analysed for their connection to social support, satisfaction with life, upward social comparison and positive or negative affect. Though it was hypothesized that using Twitter, as a text-based social media site, would not have an association with wellbeing, actively using Twitter was connected with more social support, while passively using Twitter was associated with reduced upward social comparison.

With recent developments and trends, these "timeline" features have begun to move from their initial concept. While they started out as a strictly chronological list of content from sources the user has actively chosen to subscribe to, many sites now make use of a more flexible feed of content, which may not be organised strictly chronologically, and might include "recommended" content, selected algorithmically. Algorithmically selected content may be prioritised according to the content's popularity and whether accounts connected to the user have interacted with it. [81]

Swart's paper Experiencing Algorithms [76] makes use of interviews and other qualitative research methods to explore "how young people make sense of, feel about, and engage with algorithmic news curation on social media". She identifies that a number of factors hold users back from "shaping the algorithm" of their timelines, ranging from belief their choices have no impact to the tools they use being difficult to learn.

2.3 Moderation and Timeline Management Tools

While these timeline-based sites have differing intentions and audiences, they generally aim to improve their users' experiences, offering entertainment and education, as well as being sources of information and news to many who use them [50]. Twitter has community guidelines and an in-house moderation policy which aims to detect and remove content that is deemed inappropriate for the site. Relating to

Category	Description
Violent Speech	You may not threaten, incite, glorify, or express desire for violence or harm.
Violent & Hateful Entities	You can't affiliate with or promote the activities of violent and hateful entities.
Abuse/Harassment	You may not share abusive content, engage in the targeted harassment of someone, or incite other people to do so.
Hateful conduct	You may not attack other people on the basis of race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease.

Table 2.1: Relevant categories from the Safety category of Twitter's rules. [85]

Category	Description
Platform Manipulation and Spam	You may not use Twitter's services in a manner intended to artificially amplify or suppress information or engage in behavior that manipulates or disrupts people's experience on Twitter.
Civic Integrity	You may not use Twitter's services for the purpose of manipulating or interfering in elections or other civic processes. This includes posting or sharing content that may suppress participation or mislead people about when, where, or how to participate in a civic process.
Misleading and Deceptive Identities	You may not impersonate individuals, groups, or organizations to mislead, confuse, or deceive others, nor use a fake identity in a manner that disrupts the experience of others on Twitter.
Synthetic and Manipulated Media	You may not deceptively share synthetic or manipulated media that are likely to cause harm. In addition, we may label Tweets containing synthetic and manipulated media to help people understand their authenticity and to provide additional context.

Table 2.2: Relevant categories from the Authenticity category of Twitter's rules. [85]

our definition of "disagreeable content", Twitter Rules [85] cover "safety" (see 2.1) and "authenticity" (see 2.2).

As well as the moderation of the site itself in accordance with the site's terms and conditions, means for the user to improve their experience on the site are offered in the form of timeline management tools. In the scope of this dissertation, timeline management tools are defined in a social media sense as any tool the user has at their disposal to shape the content that is presented to them on their "timeline", either through additive or subtractive curation. The typical tools common to most social media are as follows:

1. Following/Subscribing

   Where the user chooses to add a user or other source of content to the collection of sources included as part of their timeline. Depending on the site, users may be able to subscribe to organisational tags or broader "themes", meaning that they may be shown content by users they did not actively choose to follow, as long as the content fits conditions they have determined.

   Most social media involves a feature similar to this, ranging from a "friends list" to a list of subscriptions.

2. Blocking

   The subtractive opposite of subscribing, where a user can block another user or specific keyword to remove all posts from said user or mentioning the keyword in question from their timeline.

3. Reporting

   The ability to bring a post that the user believes is in violation of the site's terms of service to the attention of the site's moderation team, whether automated or manual.

4. Muting

   This feature allows a user to hide a certain user's posts and content from their timeline without unfollowing them and can be reversed at any time without informing the muted user.

2.4 Case Study: Twitter Timeline Curation

Twitter, as an active website that is constantly undergoing improvements and fixes, frequently sees its features change [61]. As a common factor brought up by study participants was frustration with recent changes to Twitter and some of these changes occurred during the span of the interview process, it is relevant to include a brief history of recent changes to the site as a case study.

2.4.1 Development of Algorithmic Timeline

Algorithmic content on Twitter began with the "best Tweets first" feature, which was introduced in February 2016 [10]. Initially, this feature had to be opted into by selecting the option in the website's settings [48], but from March 17, 2023, it became opt-out [37].

On December 18, 2018, Twitter introduced a more straightforward way to toggle between timelines - the "sparkle icon", in the top right corner of the app's homepage. This allowed users to switch between the algorithmic timeline ("Home") to the chronological one ("Latest Tweets"). Routine use of a particular setting would make it the default for that user. [49]

In March 2022, Twitter debuted an update to this setting that would allow users to view both timelines at once, on tabs that could be switched between more easily [60]. This feature was introduced on March 10 [58], but after public complaint that the new split timeline always defaulted to the algorithmic timeline, it was switched back on March 14 [75].

On January 11 2023, Twitter removed the "sparkle icon" altogether, replacing it with two tabs – For You, the algorithmic timeline, and Following, the chronological one - that the user can swipe between [20]. When it was first released, the feature would default to the For You timeline whenever it was reopened, regardless of which timeline the user had selected previously, but after widespread complaint, this was updated on February 7 [68].

2.4.2 Acquisition of Twitter by Elon Musk

Purchase

On April 14, 2022, Elon Musk made an unsolicited and non-binding offer to purchase Twitter for $43 billion [79]. Eleven days later, on April 25, Twitter's board accepted the buyout offer, with the transaction due to be completed "sometime in 2022" [72]. Musk attempted to terminate his deal to buy Twitter on July 8 [26], but became the subject of a lawsuit from Twitter on July 12 [22]. After much further controversy, Musk's legal team informed Twitter that Musk had changed his mind and offered to close the acquisition for the original agreed price [40]. The deal was closed on October 27, with Musk firing Twitter's CEO, chief financial officer (CFO) and general counsel shortly after becoming the new owner [69].

Layoffs

Musk made substantial changes to Twitter in the time immediately after acquiring the site, with substantial layoffs to the company's workforce beginning on November 4, 2022. The layoffs were described as "haphazard", with around 3700 of the company's 7500 finding that they had lost their jobs, some only finding out they were now unemployed by their company email accounts being shut off [23]. On 16 November, Musk sent employees an email asking them to "sign a pledge" and commit to "extremely hardcore" work on what he deemed "Twitter 2.0" [9] or be forced to leave with three months of severance pay [70]. In response, hundreds of employees decided to leave [34]. Between this exodus and the previous round of layoffs, many areas of Twitter were left decimated. Twitter's communications team in particular was stripped back to the point where any queries to the website's press email address received the automatic response of a single "Pile of Poo" emoji [78].

Content Moderation

On October 28, the day after he became Twitter's new owner, Musk tweeted that he planned to form a "content moderation council with widely diverse viewpoints." [46] No specific details were given for who would be on the council, and it was never actually formed, with Musk later claiming that the formation of the content moderation council was a requirement of a deal he made with "a large coalition of political/social activist groups" who later "broke the deal" [19].

In late November 2022, Musk began unsuspending banned accounts, including Jordan Peterson, The Babylon Bee and Donald Trump [13], while many other accounts were suspended at a much higher rate than before Musk's takeover [74]. The same report from the firm Bot Sentinel calculated that around 877,000 users deactivated their accounts in the time between October 27 and November 1 [74]. Twitter also began to block outgoing links to its competitor site Mastodon as "harmful" or potentially including malware [16]. A policy barring users from promoting Mastodon and other social media platforms was announced on December 18 but rescinded within a day after receiving considerable backlash [51].

Musk also made several changes to its content moderation policy. Effective from November 23 2022, Twitter stated that it would no longer be enforcing its policy on misleading information about COVID-19 [54]. On 12 December, Twitter dissolved its Trust and Safety Council, responsible for addressing hate speech, child exploitation, suicide, self-harm and other issues, shortly before it was scheduled to meet [52]. Changes were made to Twitter's hateful conduct policy, specifically, areas protecting transgender people [92], as well as its policy against violent speech [21].

In a discussion attended by representatives from several major advertisers and brands, Musk stressed that he wanted to continue battling hateful content, and emphasised the potential importance of Community Notes, a crowdsourced fact-checking feature that he suggested would "obviate the need for a lot of the content stuff that's currently in place." [28]

API changes

On January 12, 2023, Twitter cut off access to Twitter's application programming interface (API) for many third-party Twitter clients, with the explanation coming a week later on January 17 that the cutoff was "enforcing [Twitter's] long-standing API rules", though requests for clarification on which rules were broken were not answered [64]. Third-party apps were officially banned by January 19 [91]. On February 2, Twitter announced that free access to the Twitter API would no longer be supported, instead being replaced with a "paid basic tier." [90] The free API access was mostly used for entertainment bots and tools offering useful functions, and developers of non-commercial Twitter projects stated that the change would "destroy research, activism and commercial projects." [90] On February 5, Musk responded to feedback with the announcement that Twitter would enable free API for bots "providing good content that is free." [66]

Verification program

Before Musk's takeover, Twitter had a verification system, introduced in June 2009. It was intended to help distinguish legitimate users such as organisations and celebrities from parodies or impostors by distinguishing "verified" accounts with blue checkmarks next to their account names [17]. On October 30, 2022, it was reported that verified users would need to pay for a Twitter Blue [1] subscription to retain their blue checkmarks. On November 9, 2022, a premium tier was introduced that also gave users verification markers [38], but after a wave of impersonation attempts, including of Musk himself, the feature was temporarily suspended [5]. It was relaunched on December 12, with new categories of checkmarks: blue for Twitter Blue users, gold for businesses, and grey for government accounts [44].

On March 23, 2023, Twitter announced that "legacy" verification badges would be removed [73]. Though this was initially announced as beginning on April 1, no badges were removed until the announcement that the removal date would now be 20 April. After the remaining "legacy badges" were removed, a campaign was started by frustrated users to block every user with a blue checkmark – the "Block the Blue" campaign, accompanied by a dedicated account, which was suspended from the website on April 22 [1]. However, despite the apparent intention that only Twitter Blue subscribers would retain blue checkmarks, several famous users have reported having the checkmarks without having paid for or requested them [43].

---

[1]A paid subscription service for Twitter announced on June 3, 2021, which offers users additional features including an "Undo Tweet" functionality and "reader mode". [59]

2.4.3   Release of Algorithm Source Code

On Friday, 31 March 2023, Twitter released the source code for many parts of the website on Github [11]. The blog post accompanying the release stated that Twitter intended to foster "the highest possible degree of transparency", without compromising user safety and privacy or the site's ability to moderate against "bad actors" [11]. This had been a goal of Musk's from as early as the acceptance of his offer to purchase the website and was part of his plan to "increase trust" and make Twitter "better than ever" [18]. Notably, the code that powers advertisement recommendations on the site was not released.

By posting the source code on Github [2], Twitter could allow community members to make suggestions for improvements to the algorithm, with the intention to review these suggestions and incorporate some into the site's main source code [11].

The site's engineering blog went into more detail on the details of how the algorithmic timeline works [86]. It provides some statistics about what is displayed to users on the For You timeline – on average, the timeline consists of "50% In-Network tweets and 50% Out-of-Network tweets", and the ranking mechanism that decides which tweets to display takes thousands of features into account.

---

[2]"A code hosting platform for version control and collaboration," which allows users to collaborate on development projects and is frequently used for open-source projects.[29]

Chapter 3

Methods

3.1 Choice of Data Collection Method

In gathering data, my main intention was to collect data on the experiences of a sample of Twitter users on a more personal than statistical level. As I was focused on the personal experience of users, I chose to make use of qualitative data collection methods, which aim to understand a scenario or phenomenon from the viewpoint of participants [56] – "qualitative research methods are designed to help researchers understand people and the social and cultural contexts within which they live." [47] This was fitting to my intended approach, as it would allow a flexible approach to gathering data from participants and focus on the subjective experiences of Twitter users.

I will briefly survey the most commonly used qualitative data collection methods and explain the reasons they were excluded in favour of my final selection, semi-structured interviews.

Focus groups allow participants to share their experience in their own words, and are flexible, allowing follow-up questions to be asked in order to follow up lines of inquiry that emerge. Focus groups offer the benefit of gaining input from multiple perspectives, as well as encouraging participants to interact and discuss questions [56], which may offer insights that participants may not have thought of unprompted [62]. However, they do have drawbacks. Focus groups require a greater amount of planning and organisation [56], and with my participants scattered across a wide geographical range I would face the difficulty of co-ordinating any in-person or partially online meetings. As well as this, the planned topics for discussion would likely touch on topics that involve personal information, which participants might be less comfortable sharing in a more public or open discussion [71].

Focus groups, as well as interviews, have the risk of being affected by social desirability bias, where research subjects are more inclined to respond to questions with answers they believe are more socially acceptable [30]. This bias is more often encountered in situations where the participant is in the presence of another individual, even if this is only the interviewer, but can occur whenever the participant believes that they may be identified through their answers.

Methods of qualitative research that do not involve face-to-face interaction, or where the participant can remain anonymous even to the researcher, can reduce this bias [30]. Postal surveys are one such method, but participants can be less likely to go in-depth with their responses. They are also prone to the issue of lower response rates [6], which is a problem "because of the likelihood – repeatedly confirmed in practice – that people who do not return questionnaires differ from those who do." [45]

One-on-one interviews, unlike focus groups, offer a chance for the participant to discuss personal topics without an audience, and allow a rapport to be formed between the interviewer and interviewee [65]. Semi-structured interviews offered the best possibilities for following up on answers, allowing me to speak to people in depth on their experiences and follow through from initial basic details to more open-ended questions.

3.2 Interview Design

Using open-ended questions prompted the participants to share their experiences, but especially allowed me to identify the areas of Twitter that were of most concern to each participant – some singled out specific incidents they had experienced, while others had opinions on the structure and community of the site as a whole.

Name	Age	Gender	Occupation	Interview Date
Abigail	20s	Female	PhD Student (Sociology)	06/02/23
Bree	20s	Nonbinary	Student (Computer Science)	15/02/23
Charles	50s	Male	Specialist Technician	15/02/23
Daniel	60s	Male	Community Enterprise Adviser	06/02/23
Evan	20s	Male	PhD Student (Maths)	16/02/23
Francis	50s	Male	Housing Manager	17/02/23
George	20s	Nonbinary	Student (Environmental Science)	23/02/23
Haley	20s	Genderfluid	Cinema worker	24/02/23
Jess	20s	Female	Researcher (Cyber Analytics)	28/02/23
Kiera	20s	Female	Student (Computer Science)	03/03/23
Laura	40s	Female	Teacher	05/03/23
Neil	30s	Male	Researcher (Data Science)	09/03/23

Table 3.1: Participant demographic data

3.2.1 Participants

The inclusion criteria I decided on required that participants be over eighteen, currently a user of Twitter, and that they had been using Twitter for at least three months. I decided to exclude participants who had been former users of Twitter, but no longer used it at the time of the interview, in order to analyse user experience and opinions at this current moment of time.

To recruit my participants, I used a mixture of word-of-mouth and poster advertising, reaching out to my own connections on my course and more widely. I created a digital poster with the details of the study and shared it in some University groups that were likely to contain people meeting the requirements. My supervisor also shared the poster on her Twitter account, allowing us to reach people who specifically use Twitter.

Over the period of data collection, I recruited twelve participants (see 3.1), the minimum total I wished to interview based on Guest, Bunce & Johnson's [31] guideline that data saturation occurs within the first twelve interviews. Future research in this area could greatly benefit from a larger and more varied sample size, as I was limited by the time I had to recruit participants and perform interviews. The interviews took place over the span of a month, from 6 February to 9 March 2023.

3.2.2 Interview Structure

For the interviews themselves, I aimed for a rough duration for each of 45 minutes, as this was more compatible with my participants' availability and ability to focus. In practice, they varied from around 30 minutes to an hour, as interviews were allowed to come to their natural conclusion when all of the base questions had been answered and the participant had nothing more to say.

The interviews were conducted both online and in person, according to the user's preference and availability. Online interviews took place over Microsoft Teams, with both audio and video enabled to facilitate a more natural conversation. In-person interviews were conducted in private spaces on University grounds, with the interviews recorded by a device running Microsoft Teams. Teams' automatic transcription tool was used as a basis for initial transcripts of the interviews, which were then manually transcribed by myself.

The final list of guide questions asked to each participant covered:

1. Their day-to-day experience of Twitter (how long they had been using the site, how much time they spent online, and so on)

2. What content they typically see on Twitter

3. What timeline management tools they currently make use of

4. Their general user experience on Twitter, and any overarching factors that affect their enjoyment of the site

As part of the interview process, I asked participants to imagine their response to a hypothetical scenario where they were exposed to tweets that matched each of the three main categories of disagreeable content identified earlier:

1. A tweet that stated an opinion they disagreed with, worded in an argumentative or hostile way

2. A hostile tweet targeted at them with use of the tagging feature

3. A tweet containing information they knew to be misinformation or from a source they knew to be untrustworthy

While I initially used image editing software to produce falsified tweets that fit these descriptions, I found that in my pilot interview the participant became overly focused on the specific details of the tweets – commenting on the user's name and amount of likes on the post – so chose to abandon this in favour of describing what the tweet contained instead. This also allowed me to adjust the parameters of the situation and gauge if that changed the participant's response - "Would your reaction be different if it was a user with more followers? Would specific wording make you take certain action?"

3.2.3 Ethics & Accounting for the Pandemic

An ethics application for this project was submitted to the faculty research ethics committee and approved. Participants were presented with a participant information sheet and signed a consent form before each interview. All names of participants have been pseudonymised, and participants were informed that they could withdraw from the interview process at any time prior to pseudonymisation.

During the interview, I made sure to prioritise the comfort of the participants, reiterating before the interview began that they were able to pause from the interview or end it entirely if at any point they were uncomfortable or needed to take a break.

Participants were offered the choice of an online or in-person interview with no pressure to attend in-person. All interviews were performed in spaces large enough for social distancing, and I wore a face mask for the duration of face-to-face interaction.

3.3 Data Analysis

To analyse the transcripts as the main body of data collected during the interview process, thematic analysis was used. Thematic analysis is defined as "a method for identifying, analysing and reporting patterns (themes) within data"[14], and was selected as appropriate for the qualitative analysis of the transcript data in order to identify recurring themes that capture "something important about the data in relation to the research question".[14]

The codes used in analysing the data were identified both "deductively" and "inductively" – some were formulated based on the initial research question, while others were created during the coding process as recurring themes became visible.

Chapter 4

Results and Discussion

4.1 Tools and Settings for Curation

The first theme of this research concerns the specific actions taken by Twitter users to control what they see on their timeline, from the choices they make when choosing which settings to use, to the timeline management tools they do and don't make use of. The following section will discuss three themes: additive rather than negative curation, ineffectuality of reporting, and distrust of the algorithm.

Additive rather than negative curation. When asked specifically about which of the timeline curation tools they made use of, most of the participants focused on their use of the following feature, with far less mention of reporting, blocking or even muting other accounts. Kiera says that of all of Twitter's timeline curation tools, she mostly uses the "following" functionality, finding it much harder to recall instances where she used the other tools.

> "I make use of the following stuff to get more stuff I want to see, but I feel like if I there's something I don't like, I guess, when I see it, I tend to just like ignore it most of the time. . . . I guess for me at least, it feels like the, you know, following people is quite crucial to enjoying Twitter. It feels like the others, the stuff like blocking and muting, they feel less essential to using Twitter so I don't use them that much."

She described her "two reactions" to encountering disagreeable content: "either to scroll past it or engage and then become annoyed at humanity and also everything." In her view, engaging with disagreeable content often, if not always, has a negative outcome, lowering her mood until she's "decided I have no hope left for humanity", and rarely achieving anything helpful. She views scrolling past and not engaging with the disagreeable content as the "smart" choice, but admits that she does sometimes click on and read through posts even if reading them upsets her, which is the course of action most likely to lead to her using other timeline management tools:

> "I'm pretty sure I've done this thing before, where I've engaged with a bit of Twitter content I probably shouldn't have engaged with, and I become annoyed at some prick that's being misogynistic or something, and then I'll - I'm pretty sure I have before - just like reported this person, even though maybe it doesn't really count as an actual violation of the terms of service or anything, but yeah, I'm pretty sure that I've probably done that before."

Quite a few participants described how they were more likely to "just scroll past" negative content, rather than taking specific action to respond to it. This usually came with the stipulation that it would depend on the source of the post - whether they were likely to see it or similar posts again, or if it was a momentary inconvenience that likely would not reappear once they had scrolled past it.

> "I'd be more inclined to just scroll past and ignore if it's just a one-time thing." (George)

> "I tend to scroll past. It depends what kind of mood I'm in. I can be fairly mischievous when I want to be." (Francis)

> "If I didn't like it, I would ignore it, to be honest." (Abigail)

"I just ignore it, really." (Laura)

"If I don't agree with it and I don't really care about the person who has that view, then I'm going to ignore it, I guess." (Neil)

"I'd probably just ignore it to be honest." (Bree)

Taking the time to interact with a post was typically either out of some kind of responsibility (to report, reinform, etc) or of a more self-sabotaging/mischievous spirit. Most participants expressed a desire to "not get involved".

Haley speculated that users could have quite a lot of control over what content they see purely through being selective about who they follow, without needing to use other timeline curation tools.

"I'm sure if you strictly were to only use the Following page rather than the For You page, I'm sure that you'd probably see less [disagreeable content], because you can not follow the people that are trying to start arguments or being deliberately aggravating."

Ineffectuality of reporting. In general, participants did not seem to believe that their use of the Report functionality would enact any significant change to Twitter, and as such used the feature much less often. Only 6 of the participants mentioned making use of the report feature while on Twitter, and some participants expressed uncertainty whether reporting posts achieves anything, with Daniel opining that Twitter's moderation system is "pretty much non-existent."

One participant, Jess, her or that she sees as manipulative/coercive, or in instances where a user is being "dogpiled"[1]. To her, reporting inappropriate advertisements or unfair behaviour is a responsibility she has towards other users, and her way of protecting the others around her. However, she has a fatalistic attitude described how she uses the report feature quite often, on two categories of post – adverts that are not relevant to towards how much these reports actually achieve:

"I don't expect them to do anything, but if you report it, then they'll at least review it, or I think they will. I don't - again, the new Twitter, you don't know what happens behind the scenes - but I think if you just report it, then at least it's on record as this person said something."

She feels that the way reports were handled had changed during her time on Twitter, with more of a focus on "empty platitude emails" than any action being taken on the content that was reported, which is frustrating to her – "I don't want an e-mail from Twitter! I don't want to know what happens!" Rather than an actual impact on the content she wants to see less of on Twitter – for the sake of herself and other users – the response from the moderation system is more aimed at reassuring her that her report is being dealt with, but she has no way of knowing if this has an actual impact.

Jess was not the only participant to have a similar attitude regarding the likelihood that reports would have an impact on the general content of the site. Bree said that while they do occasionally report inappropriate content, they believe that that reporting is somewhat ineffectual:

"I don't really think it does have an effect, if I'm being completely honest. Twitter will send you emails being like "we've reviewed this" and "we are reviewing this" and stuff, and like "we'll let you know if anything changes", but then they don't really let you know, so it's just like.... I don't think they have a vested monetary interest in having like an effective control over hate speech, like really, let's be honest, unless it gets really egregious."

They see keeping – and encouraging – controversial content on Twitter as beneficial for the site's popularity and success, as more interaction between users, regardless of tone, is seen as "engagement", and engagement draws advertisers to the site. From this angle, reporting the content of larger and more notorious Twitter users is "sort of like shooting a water gun at the sun." Certain celebrities or public figures are seen as untouchable, due to their larger following, and any reports against them are unlikely to make a difference.

Haley also identified this relationship between how controversial content is and how much engagement or attention it gets. Their conclusion is slightly different – they don't think that Twitter is deliberately recommending controversial content, but point out a cycle that commonly occurs on the site:

---

[1]"a barrage of criticism, insults, etc. directed at someone or something by many people"[42]

> "If someone says something that could be controversial, you probably stand on one side or the other of that argument, and that deliberately entices people to interact, which means that Twitter is looking at it and going "lots of people are interacting with this, it must be popular, or at least people must want to see it". And therefore Twitter must be more likely to show it to people, would be my thoughts about it. And obviously the more people see it, the more people state their opinion, and the cycle continues, as Twitter would like it to do."

Algorithm and distrust of it. One of the most straightforward choices available to Twitter users in regards to shaping what appears on their timeline is which setting they use. The "Following" option is the one closest to the original concept of a timeline, displaying only posts from people and topics the user is following in chronological order, going father back in time the further the user scrolls down the page. The "For You" timeline option, the most recent incarnation of Twitter's experimentation with incorporating algorithmic content, was implemented in its current form in January 2023. It consists of a combination of posts from accounts the user follows, not necessarily in chronological order, as well as occasional posts from accounts not followed by the user, selected and ranked according to a number of factors including their popularity and the degree of connection they have to accounts the user does follow.

Seven of the participants said that they predominantly used the For You, algorithmically curated timeline, and four used Following. Only one participant used both timelines in equal measure, but a few said they switched between them for specific reasons.

Some of those who used the "For You" timeline were unaware that this was a setting that could be changed, or said that they simply used "the default". On the other hand, participants who had switched to the Following timeline specified their distrust or dislike of the algorithm as a key reason why. Participants tend to feel that the algorithm's recommendations are either not relevant to them or unsettlingly so – when the predictions are too accurate, they feel uncomfortable with how much Twitter seems to know about their interests.

Abigail, who discovered she was using the For You timeline after being prompted, said she would likely switch to the Following timeline - now I know that you can actually change that, I probably will change that." She expressed discomfort with the algorithm's recommendations, saying that the posts recommended are often irrelevant to her, but if they are appropriate, she "will actively be annoyed that they've got that right". The idea of a "robot" or "AI" deciding what to display to her based on her activity on the site was unsettling to her, and she expressed a need for "defiance against AI" that she achieves through not interacting with recommended content.

A number of participants do knowingly use the algorithmic timeline. Bree related that they had tried both timeline options, but ended up returning to For You as it offered more variety and content than what was provided by the Following timeline:

> "I've tried swapping it to the "in order" one, but I didn't follow enough people for it to keep my silly little brain adequately entertained. So I swapped it back. I willingly subscribed to the algorithm."

They described this as having both benefits and drawbacks – while most recommendations were interesting or appealing to them, they sometimes get recommended "stuff that... is designed to rile you up". This content was rarely recommended to them directly from the source they disagreed with, but more frequently from users whose views align with Bree "reposting" or "quote tweeting" content with a critical comment or negative lens – but, as Bree pointed out, this was still bringing the unpleasant content onto their timeline.

This issue with "quote tweeting" was also mentioned by users of the Following timeline - participant Daniel identified that people he followed would occasionally quote tweet posts from politicians he disagreed with or other sources of unpleasant content. While this was typically done with the intention to critique or mock the original tweet, Daniel expressed concern that it was actually just increasing its visibility by sharing the content into circles that would normally not interact with it at all. He pointed out that these tweets may be made deliberately controversial as "rage bait" – so that other users will have negative reactions and share them more often.

Some participants made use of both timelines, for different reasons – Evan, who mostly views Twitter as a means of entertainment, said that he generally used Following, but, like Bree, swapped over to For You when he was particularly bored:

> "I'd say, really the only time I go for For You is literally "I'm on Twitter because I'm bored, and I'm so bored that I've cycled back through all the other apps I could be scrolling through.

> I'm now scrolling through Twitter and I've run out of things on Following, so now I'll scroll For You". But that doesn't happen very often."

He viewed the For You page as a supplement to the content he usually views – "scraping the barrel" – as the algorithm only sometimes provides him with content he is interested in, though he did comment that sometimes the recommendations are "worryingly good", saying that Twitter's decisions on which content he would be interested in can be "eerie" when they are right.

The lack of transparency on what drives the algorithm's decisions is another factor that the participants identified as making them mistrust recommended content. Few participants identified which factors they thought affected the algorithm's decisions, and a number of them expressed concern about what the algorithm, and by extension Twitter, knew about them. Jess commented:

> "The reason that I don't want the recommended content is because I don't know what Twitter has decided that they know about me. . . I don't know how these algorithms work. There's very little transparency."

Participants who prefer the chronological timeline identified its main appeal as an ability to have more control over what their timeline contained – by choosing who they follow and unfollow, they decide what content appears on their feed. Evan describes browsing the For You page as "opening the floodgates" in comparison to the more cultivated content he gets on the Following timeline: "at least when I restrict the number of accounts I follow, then I can kind of have some sort of agency of how much clutter I get."

Chapter 5

Critical Evaluation

5.1 Overall Process

Over the lifespan of the project, from initial selection of the topic and research questions to concluding analysis on my data, I have learned a lot – not just about the subject of Twitter, but about qualitative data methods and how to plan and execute a large-scale project like this dissertation.

I particularly enjoyed the process of interviewing participants and exploring their experiences – I found I learned more than expected from each interview. While the initial topic guide set a structure to ensure I could compare discussed topics across interviews, it was when participants went "off script" and began to talk about other aspects of their experience on Twitter that I gained deeper insight into what it was they needed from the website and its community as a whole.

Along similar lines, there were a number of themes identified from my data that were a surprise to me. When initially considering the areas my research would look into, I did not expect the choice of algorithmic or non-algorithmic timeline to have anywhere near as significant an impact as other timeline management tools, and included a question about it as one of my "warm-up" questions near the beginning of the interview. However, it emerged that many of my participants had strong opinions about the algorithm and its effect on their Twitter experience, and it became one of the most prevalent themes in my analysis.

5.2 Drawbacks of Data Collection

My data collection process could have been considerably more vigorous – I only interviewed twelve participants over the duration of the project, and believe that a larger sample size would have gathered a broader range of experiences and lent more confidence to the trends identified. I think it likely that the methods chosen to spread word of the study were not suited to finding participants who fit the inclusion criteria – most of those who I mentioned the study to in passing said that they did not use Twitter, leading me to suspect that I did not manage to make contact with the right audience of Twitter users. If the recruitment process had relied less on word-of-mouth and posters shared among my own peers, instead putting more focus on spreading word on Twitter itself, I believe I would have been able to recruit more participants who met the inclusion criteria, as a large proportion of those who I did go on to interview discovered my study through the sharing of the digital poster on Twitter.

Due to the limited number of participants, it is hard to know whether the trends identified among my subjects are generalisable or merely due to the demographics I interviewed. Many of my participants reported a generally "below the radar" experience with Twitter – "rarely does anybody feel that it's important enough that they comment about me" but some did specifically identify that they thought this was due to not being a member of any marginalised groups, who might have different experiences with online harassment and other disagreeable content.

Over the time period from recruiting participants to beginning thematic analysis, a disproportionate amount of time was spent manually correcting the automatically generated interview transcripts. The amount of time spent on ensuring all transcripts were entirely correct in grammar and spelling was entirely disproportionate to its necessity, and it would have been considerably more time-efficient to only apply basic correction to the automatic transcriptions and only fully proofread them if they were required for quotes in the analysis. Another possible alternative could have been to make use of a professional transcription service to save time and ensure the completed transcripts were good quality, though this

would require more consideration of data protection and likely an extension to the ethics application, as the users' data would no longer be being handled exclusively by myself.

5.3   Future Development

In order to further investigate the themes identified in my study, I recommend that future study involve a far wider range of participants, particularly from demographics not covered by my existing research. A more purposeful sampling method, selecting participants based on their experience with a given phenomenon, may allow the selection of participants who have had experience with certain circumstances on Twitter which might affect their user experience, such as members of marginalised groups who are more vulnerable to harassment and other disagreeable content.

I personally believe that my insights could be greatly improved by the opportunity to interview more participants from a wide variety of backgrounds. With Twitter's current situation remaining extremely fast-changing, it is vitally important to continue to monitor the experience of its users over time. While capturing users' reactions to the changes to date has provided valuable insight, it is by no means a "one-and-done" investigation - the environment, and, in turn, the user experience, is constantly changing.

Chapter 6

Conclusion

This paper set out to explore the experiences of Twitter users regarding the timeline curation tools they use and how they could be achieved. Through an interdisciplinary approach, both analysing the existing tools and interviewing users to identify their key concerns and patterns of use, I have been able to answer my original research questions:

1. How do users make use of Twitter's existing timeline curation tools, and do these tools fit the users' needs?

Interviews with twelve existing Twitter users revealed that participants often do not utilise the full range of timeline curation tools available to them, though some of the issues they identified would be solved by making use of tools they were not aware of or did not have the motivation to begin using.

2. How does Twitter's algorithmic timeline affect the quality of experience for users?

The interviews also showed that users typically fell into one of two groups when it came to use of the algorithmic timeline – either greatly suspicious of it due to its lack of transparency, or willing to accept it as the "default" as long as it continued to provide them with new, appropriate content.

Based on these discoveries, I recommend several means of improving the experience for users of Twitter's timeline feature:

- Provide users with better information on the tools available to them, especially when these tools change

- Increase transparency on the factors affecting what users see and the results of actions they take (reporting, etc.)

- Improve moderation of the site, taking action more quickly against misinformation and reported content, to reduce the responsibility of removing disagreeable content for users

Future research should aim to survey a wider range of participants and identify ways these changes could be implemented. These changes could not only improve user experience on Twitter, but help to improve moderation and decrease disagreeable content on social media sites in general, making them a safer and more beneficial space for social interaction, communication and news circulation on the Internet.

Bibliography

[1] Kelsey Ables. Twitter users #blocktheblue as 'verified' accounts take on new meaning. *The Washington Post*, 2023.

[2] Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2):211–236, 2017.

[3] Jack Bandy and Nicholas Diakopoulos. Curating quality? How Twitter's timeline algorithm treats different types of news. *Social Media + Society*, 7(3):20563051211041648, 2021. `arXiv:https://doi.org/10.1177/20563051211041648`, `doi:10.1177/20563051211041648`.

[4] Sarah Aragon Bartsch. Digital wellbeing in the context of decision support systems. *Position Paper*, 2019.

[5] Ashley Belanger. Twitter quietly drops $8 paid verification; "tricking people not ok," musk says. *Ars Technica*, 2022.

[6] Judith Bell and Stephen Waters. *Ebook: doing your research project: a guide for first-time researchers*. McGraw-hill education (UK), 2018.

[7] Jonathan Bishop. The effect of de-individuation of the internet troller on criminal procedure implementation: An interview with a hater. *International journal of cyber criminology*, 7(1), 2013.

[8] Jonathan Bishop. Representations of 'trolls' in mass media communication: a review of media-texts and moral panics relating to 'internet trolling'. *International Journal of Web Based Communities*, 10(1):7–24, 2014.

[9] Twitter Blog. Twitter 2.0: Our continued commitment to the public conversation, 2022. URL: `https://blog.twitter.com/en_us/topics/company/2022/twitter-2-0-our-continued-commitment-to-the-public-conversation`.

[10] Twitter Blog. Never miss important tweets from people you follow, 2023. URL: `https://blog.twitter.com/official/en_us/a/2016/never-miss-important-tweets-from-people-you-follow.html`.

[11] Twitter Blog. A new era of transparency for twitter, 2023. URL: `https://blog.twitter.com/en_us/topics/company/2023/a-new-era-of-transparency-for-twitter`.

[12] Leticia Bode and Emily K Vraga. In related news, that was wrong: The correction of misinformation through related stories functionality in social media. *Journal of Communication*, 65(4):619–638, 2015.

[13] Russell Brandon. Elon musk begins reinstating banned twitter accounts, starting with jordan peterson and the babylon bee. *The Verge*, 2022.

[14] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006. URL: `https://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa`, `arXiv:https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa`, `doi:10.1191/1478088706qp063oa`.

[15] Christopher Burr, Mariarosaria Taddeo, and Luciano Floridi. The ethics of digital well-being: A thematic review. *Science and engineering ethics*, 26(4):2313–2343, 2020.

[16] Peter Butler. What is mastodon, the alternative social network now blocked by twitter? *CNET*, 2022.

[17] Pete Cashmore. Twitter launches verified accounts. *Mashable*, 2009.

[18] Cision. Elon musk to acquire twitter. *PR Newswire*, 2023.

[19] Mitchell Clark. Elon musk tries to blame 'activists' for his twitter moderation council lie. *The Verge*, 2022.

[20] Mitchell Clark. Twitter defaults to a for you page now, just like tiktok. *The Verge*, 2023.

[21] Mitchell Clark. Twitter rewrites its rules on violent content under elon musk. *The Verge*, 2023.

[22] Kate Conger and Lauren Hirsch. Twitter sues musk after he tries backing out of $44 billion deal. *The New York Times*, 2022.

[23] Kate Conger, Ryan Mac, and Mike Isaac. Confusion and frustration reign as elon musk cuts half of twitter's staff. *The New York Times*, 2022.

[24] Naomi Craker and Evita March. The dark side of facebook®: The dark tetrad, negative social potency, and trolling behaviours. *Personality and Individual Differences*, 102:79–84, 2016.

[25] Dictionary.com. "misinformation". https://www.dictionary.com/browse/misinformation, 2023. Accessed: 2023-04-28.

[26] Clare Duffy, Brian Fung, and Rachel Metz. Musk tells twitter he wants out of deal to buy it. twitter says it will force him to close the sale. *CNN*, 2022.

[27] Maeve Duggan. Part 1: Experiencing online harassment, Dec 2019. URL: https://www.pewresearch.org/internet/2014/10/22/part-1-experiencing-online-harassment.

[28] Brian Fung. Elon musk pleads with advertisers to stay on twitter. *CNN Business*, 2022.

[29] Github. Get started. *Github Docs*, 2023.

[30] Pamela Grimm. Social desirability bias. *Wiley international encyclopedia of marketing*, 2010.

[31] Greg Guest, Arwen Bunce, and Laura Johnson. How many interviews are enough? an experiment with data saturation and variability. *Field methods*, 18(1):59–82, 2006.

[32] Jason Hannan. Trolling ourselves to death? social media and post-truth politics. *European Journal of Communication*, 33(2):214–226, 2018.

[33] Lars Kai Hansen, Adam Arvidsson, Finn Årup Nielsen, Elanor Colleoni, and Michael Etter. Good friends, bad news-affect and virality in twitter. In *Future Information Technology: 6th International Conference, FutureTech 2011, Loutraki, Greece, June 28-30, 2011, Proceedings, Part II*, pages 34–43. Springer, 2011.

[34] Alex Heath and Mia Sato. Hundreds of employees say no to being part of elon musk's 'extremely hardcore' twitter. *The Verge*, 2022.

[35] Alex Hern. Twitter CEO: We suck at dealing with trolls and abuse. *The Guardian*, 5, 2015.

[36] Shagun Jhaver, Larry Chan, and Amy Bruckman. The view from the other side: The border between controversial speech and harassment on kotaku in action. *arXiv preprint arXiv:1712.05851*, 2017.

[37] Matt Kamen. Twitter's non-chronological timeline is now opt-out. *Wired Magazine*, 2016.

[38] Jennifer Korn and Donie O'Sullivan. Elon musk delays $8 'blue check' twitter verification plan until after the midterms. *CNN Business*, 2022.

[39] Yabing Liu, Chloe Kliman-Silver, and Alan Mislove. The tweets they are a-changin': Evolution of twitter users and behavior. In *Proceedings of the international AAAI conference on web and social media*, volume 8, pages 305–314, 2014.

[40] Cara Lombardo and Dana Cimilluca. Elon musk proposes closing twitter deal on original terms. *The Wall Street Journal*, 2022.

[41] Alexandra Masciantonio, David Bourguignon, Pierre Bouchat, Manon Balty, and Bernard Rimé. Don't put all social network sites in one basket: Facebook, Instagram, Twitter, TikTok, and their relations with well-being during the COVID-19 pandemic. *PLOS ONE*, 16(3):1–14, 03 2021. doi: 10.1371/journal.pone.0248384.

[42] Merriam-Webster. "dogpile". https://www.merriam-webster.com/dictionary/dogpile, 2023. Accessed: 2023-04-28.

[43] Steve Mollman. Elon musk weighs in on campaign to block twitter accounts that still have blue check mark after legacy purge. *Fortune*, 2023.

[44] Sara Morrison. Twitter enters its chaotic new multicolored, multishaped check mark phase. *Vox*, 2022.

[45] Claus Adolf Moser and Graham Kalton. *Survey methods in social investigation*. Routledge, 2017.

[46] Elon Musk. "twitter will be forming a content moderation council with widely diverse viewpoints.", 2022. URL: https://twitter.com/elonmusk/status/1586059953311137792.

[47] Michael D Myers and David Avison. *Qualitative research in information systems: a reader*. Sage, 2002.

[48] Casey Newton. Twitter begins rolling out its algorithmic timeline around the world. *The Verge*, 2016.

[49] Casey Newton. Twitter is relaunching the reverse-chronological feed as an option for all users starting today. *The Verge*, 2018.

[50] Rasmus Kleis Nielsen and Kim Christian Schrøder. The relative importance of social media for accessing, finding, and engaging with news: An eight-country cross-media comparison. *Digital journalism*, 2(4):472–489, 2014.

[51] Matt O'Brien and Brian P. D. Hannon. Twitter deletes new policy banning 'free promotion' of rival social media platforms. *USA Today*, 2022.

[52] Matt O'Brien and Barbara Ortutay. Musk's twitter disbands its trust and safety advisory group. *AP News*, 2022.

[53] Jonathan Oliver, Paul Pajares, Christopher Ke, Chao Chen, and Yang Xiang. An in-depth analysis of abuse on twitter. *Trend Micro*, 225:1–22, 2014.

[54] Donie O'Sullivan. Twitter is no longer enforcing its covid misinformation policy. *CNN Business*, 2022.

[55] Brian L Ott. The age of twitter: Donald j. trump and the politics of debasement. *Critical studies in media communication*, 34(1):59–68, 2017.

[56] Cathryne Palmer and Amanda Bolderston. A brief introduction to qualitative research. *Canadian Journal of Medical Radiation Technology*, 37(1):16–19, 2006.

[57] Jessica A. Pater, Moon K. Kim, Elizabeth D. Mynatt, and Casey Fiesler. Characterizations of online harassment: Comparing policies across social media platforms. In *Proceedings of the 2016 ACM International Conference on Supporting Group Work*, GROUP '16, page 369–374, New York, NY, USA, 2016. Association for Computing Machinery. doi:10.1145/2957276.2957297.

[58] Sarah Perez. Twitter's latest update makes it easier to escape its recommendation algorithm. *TechCrunch*, 2022.

[59] Sarah Perez. Twitter launches its premium subscription, twitter blue, initially in canada and australia. *TechCrunch*, 2023.

[60] Jay Peters. Twitter makes it harder to choose the old reverse-chronological feed. *The Verge*, 2022.

[61] Twitter Developer Platform. Changelog, 2023. URL: https://developer.twitter.com/en/updates/changelog.

[62] Stephen Polgar. *Introduction to research in the health sciences / Stephen Polgar, Shane A. Thomas.* Churchill Livingstone/Elsevier, Edinburgh, sixth edition. edition, 2013.

[63] Cambridge University Press. "harassment". `https://dictionary.cambridge.org/dictionary/english/harassment`, 2023. Accessed: 2023-05-04.

[64] Kevin Purdy. Twitter says third-party apps broke "long-standing api rules," won't name rules. *Ars Technica*, 2023.

[65] Sandy Q Qu and John Dumay. The qualitative research interview. *Qualitative research in accounting & management*, 8(3):238–264, 2011.

[66] Emma Roth. Elon musk says bots with 'good content' can use twitter's api for free. *The Verge*, 2023.

[67] Florian Saurwein and Charlotte Spencer-Smith. Automated trouble: The role of algorithmic selection in harms on social media platforms. *Media and Communication*, 9(4):222–233, 2021.

[68] Stan Schroeder. Twitter stops forcing 'for you' timeline on ios and android users. *Mashable*, 2023.

[69] Faiz Siddiqui and Elizabeth Dwoskin. Elon musk acquires twitter and fires top executives. *The Washington Post*, 2022.

[70] Faiz Siddiqui and Jeremy B. Merrill. Musk issues ultimatum to staff: Commit to 'hardcore' twitter or take severance. *The Washington Post*, 2022.

[71] Janet Smithson. Focus groups. *The Sage handbook of social research methods*, 357:370, 2008.

[72] Todd Spangler. Elon musk clinches deal to buy twitter for $44 billion. *Variety*, 2022.

[73] Todd Spangler. Twitter to revoke 'legacy' verified badges in april, leaving only paying subscribers with blue check-marks. *Variety*, 2023.

[74] Chris Stokel-Walker. Twitter may have lost more than a million users since elon musk took over. *MIT Technology Review*, 2022.

[75] Twitter Support. We heard you — some of you always want to see latest tweets first., 2022. URL: `https://twitter.com/TwitterSupport/status/1503443926258180106`.

[76] Joëlle Swart. Experiencing algorithms: How young people understand, feel about, and engage with algorithmic news selection on social media. *Social Media + Society*, 7(2):20563051211008828, 2021. `arXiv:https://doi.org/10.1177/20563051211008828`, `doi:10.1177/20563051211008828`.

[77] Thi Tran, Rohit Valecha, Paul Rad, and H Raghav Rao. An investigation of misinformation harms related to social media during humanitarian crises. In *Secure Knowledge Management In Artificial Intelligence Era: 8th International Conference, SKM 2019, Goa, India, December 21–22, 2019, Proceedings 8*, pages 167–181. Springer, 2020.

[78] Rachel Treisman. Got a question for twitter's press team? the answer will be a poop emoji. *NPR*, 2023.

[79] Giles Turner and Maxwell Adler. Elon musk makes $43 billion unsolicited bid to take twitter private. *Bloomberg*, 2022.

[80] Twitter. What is the twitter follow limit and ratio?, Mar 2019. URL: `https://help.twitter.com/en/using-twitter/twitter-follow-limit`.

[81] Twitter. About your for you timeline on twitter, 2023. URL: `https://help.twitter.com/en/using-twitter/twitter-timeline`.

[82] Twitter. How to block accounts on twitter, 2023. URL: `https://help.twitter.com/en/using-twitter/blocking-and-unblocking-accounts`.

[83] Twitter. How to mute accounts on twitter, 2023. URL: `https://help.twitter.com/en/using-twitter/twitter-mute`.

[84] Twitter. How to mute accounts on twitter, 2023. URL: https://help.twitter.com/en/safety-and-security/report-abusive-behavior.

[85] Twitter. The twitter rules: safety, privacy, authenticity and more, 2023. URL: https://help.twitter.com/en/rules-and-policies/twitter-rules.

[86] Twitter. Twitter's recommendation algorithm, 2023. URL: https://blog.twitter.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm.

[87] Tom Van Laer. The means to justify the end: Combating cyber harassment in social media. *Journal of Business Ethics*, 123(1):85–98, 2014.

[88] Mariek M P Vanden Abeele. Digital Wellbeing as a Dynamic Construct. *Communication Theory*, 31(4):932–955, 10 2020. arXiv:https://academic.oup.com/ct/article-pdf/31/4/932/41146766/qtaa024.pdf, doi:10.1093/ct/qtaa024.

[89] Yuxi Wang, Martin McKee, Aleksandra Torbica, and David Stuckler. Systematic literature review on the spread of health-related misinformation on social media. *Social science & medicine*, 240:112552, 2019.

[90] Jess Weatherbed. Twitter replaces its free api with a paid tier in quest to make more money. *The Verge*, 2023.

[91] Kyle Wiggers. Twitter officially bans third-party clients after cutting off prominent devs. *TechCrunch*, 2023.

[92] Angela Yang. Twitter quietly changes its hateful conduct policy to remove standing protections for its transgender users. *NBC News*, 2023.

Appendix A

Interview Topic Guide

INTERVIEW TOPIC GUIDE

1. How long have you been using Twitter?
2. Can you describe how you use Twitter throughout the day? What timeline settings do you use?
3. What kind of content do you typically see on Twitter?
4. Tell me about the last time you saw content on Twitter that you didn't like.
5. What's your normal reaction when encountering this kind of content?

The interviewer will define what timeline management tools are for the interviewee.
Timeline management tools include:
- *Block feature*
  - o *Allows a user to "block" another user, restricting that user from contacting them, seeing their posts, or following them.*
- *Subscribe feature*
  - o *Allows the user to subscribe to another user's posts, meaning that content posted by that user always appears on their timeline.*
- *Report feature*
  - o *Brings a post to the attention of the website's manual or automated moderation process, in order for it to be removed from the site if it is deemed in violation of the site's content policy. Typically, also removes the post from the user's timeline.*

6. Can you tell me about the last time you used one of Twitter's timeline management tools?

The interviewer will show the interviewee a number (dependent on time) of "mockup" Twitter screenshots, showing a range of falsified posts featuring "unpleasant" content. (See below).

7. If you saw this post on your Timeline and wanted to prevent seeing any more like it, how would you go about it?

8. Who do you see content from most frequently on your Timeline?
9. When was the last time you remember Twitter recommending content to you? What was your opinion on the content?

10. Considering your experiences with Twitter's timeline management tools, do you have any suggestions for ways they could be improved? Any additional tools you think would be helpful?

Appendix B

Interview Participant Consent Form

Martha Baylis

University of BRISTOL

CONSENT FORM

Timeline curation: Safety and wellbeing practices of Twitter users

Please answer the following questions to the best of your knowledge.	YES	NO
HAVE YOU:		
been given information explaining the study?	O	O
had an opportunity to ask questions and discuss the study?	O	O
received satisfactory answers to any questions you asked?	O	O
received enough information about the study for you to decide about your participation?	O	O
DO YOU UNDERSTAND:		
that you are free to withdraw from the study at any time?	O	O
that you may withdraw without having to give any reason?	O	O
that you are free to withdraw your data at any point prior to the point of anonymisation? (one week following data collection)	O	O
DO YOU CONSENT:		
to the conversation being audio-recorded for the purpose of transcription?	O	O

I hereby fully and freely consent to my participation in this study.

Participant's signature: _____   Date: _____

Name in block letters: _____

If you have any concerns related to your participation in this study, please direct them to the Faculty of Engineering Research Ethics Committee, via the Research Governance Team ; research-governance@bristol.ac.uk

Appendix C

Codebook

C.1 Additive rather than subtractive curation

C.1.1 Scrolling past

Description: Participant describes scrolling past disagreeable content.

Example: "I'd be more inclined to just scroll past and ignore if it's just a one-time thing."

C.1.2 Not using timeline curation tools

Description: Participant mentions a specific timeline curation tool they don't make use of or use less frequently.

Example: "I don't tend to block people that much."

C.2 Ineffectuality of reporting

C.2.1 Ineffectuality of reporting

Description: Participant expresses a belief that reporting content on Twitter will have little to no effect.

Example: "I just feel like reporting is sort of like shooting a water gun at the sun."

C.2.2 Controversy good for Twitter

Description: Participant states a belief that Twitter profits from having disagreeable content on their site.

Example: "If you report something, I don't think it's in their best interest to cut off that person because they wanna cultivate sort of a space that has this kind of discourse, where people come on and they have arguments, because it creates engagement and they wanna sell that engagement to their advertisers."

C.3 Distrust of the Algorithm

C.3.1 Choice of timeline

Description: Participant states which of Twitter's timeline settings they use.

Example: "I use Following as opposed to For You."

C.3.2 Distrust of the algorithm

Description: Participant states discomfort or distrust in what Twitter's algorithm shows them.

Example: "I think one thing that sort of in some ways puts me off is sometimes it's worryingly good. I think that's something that's a bit off putting."