

History and Contrarian Expression: Debating Genocide on Reddit

Social Media + Society
October-December 2025: 1–12
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20563051251382449
journals.sagepub.com/home/sms



Aliaksandr Herasimenka^{1,2}  and Ralph Schroeder¹ 

Abstract

Social media communities are increasingly considered as spaces where evidence-based information about history is undermined. We combine expert interviews and content analysis of debates about genocide in a popular online contrarian community to understand how to mitigate the spread of misleading information across such communities in the domain of history. We analyse 1725 entries on 10 Reddit forums dedicated to debating and promoting scepticism towards international consensus about prominent historical topics. The entries we analyse cover three topics referred to as genocide by forum members: the Holocaust, the Holodomor, and the COVID-19 vaccination. The prevailing view suggests that contrarian expression fosters smaller, ideologically homogeneous, and relatively radical online spaces where differing views are diminished or entirely absent. To the contrary, we analyse real-life behavioural data to demonstrate substantial scepticism towards contrarian narratives even in some of the most popular dedicated online communities, which may suggest that these spaces are more internally contested than prevailing theories imply.

Keywords

contrarian expression, history, genocide, disinformation, Reddit

Introduction

The spread of misleading information¹ about historical events increasingly puts social media in the spotlight as spaces where verified information about issues of common concern is undermined (Jungheer & Schroeder, 2021). This is particularly urgent with respect to contested historical events. Global emergencies like the COVID-19 pandemic (Sturm & Albrecht, 2021) and the context of countries like China that lack an open internet (Farrell, 2016) are bound to intensify the contestation around the interpretation of ongoing events that will pass into history.

Debates and disagreements are an essential part of the public sphere. However, some online communities tend to host more misleading information than others and are frequently dedicated to debating and promoting statements that are sceptical of an international consensus about historical events (we label such statements ‘contrarian expression’ and their spaces ‘contrarian communities’; see Supplemental Appendix B for the full definitions). We aim to understand how to mitigate the spread of misleading information across such communities in the domain of history.

We fulfil this aim by addressing two objectives. First, to build an empirical foundation for considering mitigating

solutions, we analyse engagement with user-generated contrarian expression. This analysis is a necessary prerequisite for normative recommendations, as the literature on mitigating misleading information in other domains, beyond history, commonly debates platform engagement mechanisms as a key immediate solution to the problem, but offers conflicting views on their effects (International Panel on the Information Environment, 2023). We use content analysis and regression analysis to *examine the link between platform engagement mechanisms* (endorsement and replies) and *contrarian expression that is sceptical or manipulative in relation to facts about genocide*. This provides evidence to inform whether platforms should, for instance, remove such expression or foster counter-engagement. Second, to address the normative question more directly, we interview experts to *evaluate alternative solutions* to reduce the spread of misleading information across online contrarian communities.

¹University of Oxford, UK

²University of Liverpool, UK

Corresponding Author:

Ralph Schroeder, Oxford Internet Institute, University of Oxford, Oxford OX2 6GG, UK.

Email: ralph.schroeder@oii.ox.ac.uk



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License

(<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

While debates about misleading information online have come to the foreground in areas like political extremism or vaccine information (e.g., Bright et al., 2022), in the literature to date, there has been less discussion of online platform-based user-generated misleading information about history. We rely on real-life behavioural data originating from online debates that contain contrarian expression on three topics referred to as genocide by online community members: the Holocaust, the Holodomor, also known as the Great Ukrainian Famine of 1932–33, that killed as many as four million people under Stalin, and COVID-19 vaccination programmes that contrarian groups often describe as ‘an act of genocide’ (see more about these cases in Methods and Cases). We focus on two types of contrarian expression: one that puts forward a position that is sceptical of the international consensus about a topic (a contrarian statement) and one that manipulatively questions the foundations of the fact-based interpretation of these events though not overtly sceptical (distortion). We also analyse expert interviews to contextualise our findings from the perspective of the extensive discussions about the objectivity of history in relation to online debates.

The Holocaust is often considered a key reference in the literature on online communication about genocide. Distortion of the Holocaust is far more common online than denying it outright (Bright 2022), specifically when a distortion falsely equates it with other historical or contemporary events, such as COVID-19 vaccination policies. This suggests that outright falsehoods about history may be less dangerous online than a deliberate distortion in the representation of historical events that suits contemporary agendas, as with conspiracy theories relating to current politics.

Throughout history, genocide has been a charged topic in which contrarian views have been especially heated. In online spaces, there is often a possibility that contributors to debates join in to agree with each other, leading to a ‘perceived consensus’ affecting opinion formation (Lewandowsky et al., 2019). Researchers have described such communities as ‘filter bubbles’ or catalysts for extremist radicalisation (Wojcieszak, 2010). To counter this trend, research has investigated promoting engagement with cross-cutting views and opinions (Elgesem et al., 2015; Oswald & Bright, 2022).

This study builds on previous literature, discussed extensively in the next section, about the spread of misleading information and contrarian expression online, which have informed our hypotheses (Krafft & Donovan, 2020; Lewandowsky et al., 2019; Oswald & Bright, 2022; Robison, 2020). To anticipate, our study reveals substantial scepticism towards contrarian expression across history-focused debates, which may challenge the notion that such views foster smaller, ideologically homogeneous, and relatively radical online spaces where differing views are diminished or entirely absent. Expert interview brought a humanities perspective on the mitigating solutions debate, emphasising that

the mitigation of misleading information should consider a non-technical perspective, bring people to work with communities, and address the gatekeeping functions of platforms and their economic model. This finding can support efforts to counter the spread of misleading information in contrarian communities.

Theory and Hypotheses

Theories about the communication of controversial topics commonly focus on how people respond to statements that challenge people’s beliefs (Krafft & Donovan, 2020; Lewandowsky et al., 2019; Robison, 2020). These theories are often built on data from experiments that draw highly conflicting conclusions. While it might seem intuitive that exposure to opposing viewpoints can lead to some degree of agreement, Oswald and Bright (2022, see also Schroeder 2022) find that this exposure online leads to a hardening of extreme positions. Further empirical evidence demonstrates that challenging ideas that contravene an individual’s established views can cause a feeling of dissonance (Harmon-Jones & Mills, 1999).

We analyse community data on debates about controversial topics collected from Reddit, a popular platform for discussing current affairs. Reddit affords forum-like debates and user-driven content ranking, making it a useful object for studying online discussions. It was the eleventh most visited website in the world and the sixth most visited in the United States as of November 2021 (Amaya et al., 2021). Reddit users, who are normally pseudonymous, contribute with threads (or discussion boards) and comments across clearly defined dedicated communities known as ‘subreddits’, where they discuss specific topics, such as football or hobbies. Everyone can join the discussion, though some are only open to users above 18 years old. Comments can take the form of text, images, videos, or links to other websites. Authorised members moderate comments, and they have the power to remove content or users. Reddit is a good case for studying contrarian expression in online communities because its subreddits house numerous contrarian groups and the platform encourages users to rate others’ comments by up- or down-voting, which can help gauge if a comment is supported or rejected by the community. We also refer to this voting as expression reinforcement.

Multiple design features can potentially influence contrarian expression on platforms like Reddit. These include platform engagement mechanisms (Neff et al., 2012), the use of pseudonyms (Herasimenka, 2022), and moderation policies. The latter issue is specifically important for the forum-like Reddit discussion space. In communities where users discuss history, the authority that subreddit moderators exercise helps them to encourage or prevent debates, while different moderation styles can lead to more participation and proper sourcing of materials (Linkevicius de Andrade & Vasques Filho, 2022).

However, as elsewhere, the strictness of community moderation on Reddit varies. A commonly posed explanation for the spread of contrarian expression is that less-moderated communities are relatively closed and ideologically homogeneous spaces cutoff from diverse viewpoints (Herasimenka et al., 2023; Marwick & Lewis, 2017). We investigate these communities by focusing on the type of engagement and the frequency of expression by their members: the score (the balance of up- and down-votes) each comment on the platform receives, the number of replies to it, as well as the frequency with which a given user posts comments in a community.

First, each comment on Reddit receives a rank of scores calculated as a sum of up- and down-votes by users. These scores affect how content is displayed in the feeds of users. Theories in information studies and social psychology argue that perceived threats or criticism from outgroup members may lead community members to rate ingroup-provided content more favourably (Chinn et al., 2023; Jost & Amodio, 2012). Hence, groups like r/conspiracy, where contrarian membership is prominent, can see contrarian expression reinforced through up-voting behaviour. Moreover, users can manipulate Reddit scores to reduce their cognitive dissonance from opposing views (Harmon-Jones & Mills, 1999). Our hypothesis thus posits that:

Hypothesis 1 (H1): *Contrarian expression may be reinforced via up-voting, while other types of expression are devalued.*

Second, does contrarian expression trigger a higher level of activity within a community? Evidence suggests that threatened ingroup members seek to affirm identities by persuading others of their position and positive image (Klein et al., 2007), with online contrarian argument potentially helping in this affirmation. These actions might be especially prevalent among ingroup members who believe an audience is observing their performance, especially when their status is uncertain (Klein et al., 2007). On Reddit, contradicting a view may be signalled by commenting, with contrarian expression signalling threatening and motivating stronger responses. This leads us to the second hypothesis:

Hypothesis 2 (H2): *Contrarian expression will trigger a more active discussion through comment replies than non-contrarian expression.*

Our final hypothesis shifts the focus to the audiences of contrarians: how fast and extensively does contrarian expression reach wider audiences? Can engagement with contrarian expression trigger a higher volume of similar expression immediately, so that manipulative content spreads faster? Selection effects make it difficult to predict their subsequent actions. Unsuccessful coping with cognitive dissonance could result in disengagement, where individuals cease community

engagement to avoid dissonance-inducing information (Case et al., 2005). However, audiences can continue to consume contrarian content as a component of meaning-making and identify reaffirmation through their engagement, even after being exposed to evidence and facts (Marwick, 2018). Assuming contrarian groups' expression can reinforce contrarian views, we propose our third hypothesis:

Hypothesis 3 (H3): *Engagement with contrarian expression makes users more likely to continue their active participation in the community discussion.*

Methods and Cases

Online Communities: Case Selection and Background

To sample the cases of the discussions of controversial historical events, we focused on the theme of genocide. We selected this theme after an initial screening of the r/conspiracy community on Reddit. This is perhaps one of the largest online communities in English dedicated to debating and promoting scepticism towards international consensus about key historical issues. It had almost two million members who 'challenge issues which have captured the public's imagination, from JFK and UFOs to 9/11', says the community's self-description (<https://www.reddit.com/r/conspiracy>). The screening showed that some of the most commented threads on this subreddit related to history contained 'genocide' in the headline. Indeed, it is arguably among the most debated topics in recent history and consistently widely debated on Reddit. We should add parenthetically that the data analysis preceded the recent hostilities in the Gaza Strip and Israel, which started on 7 October 2023, in relation to which the term genocide has again been heavily contested. While this conflict falls outside the scope of our study, this context nevertheless points to the relevance of our analysis, which we address by discussing how future work might build on our study.

We used the Reddit functionality to sample the threads mentioning the theme of genocide in the title by the following criteria: the amount of engagement in the form of comments they received since the establishment of the r/conspiracy community and the level of controversy as marked by the platform algorithms based on users' ratio of up- and down-voting in the community. This sampling strategy resulted in three sets of threads covering three distinct cases related to genocide that attracted substantial attention among users: the Holocaust, the Holodomor, and COVID-19 vaccination as 'genocide'. The Holodomor preceded the Holocaust, whereas COVID-19 is closer to the present, thus providing a wide range. See the metadata summary for the sampled threads in Supplemental Appendix A. One of the collected threads on the Holodomor is from a separate community called r/Documentaries, where users discussed documentary films.

This specific thread was extensively referred to in *r/conspiracy*, so we included this thread in our sample. We anonymised this publicly available data and removed all metadata that might identify its authors. The data-collection procedures followed the platform's terms and conditions, utilised an official API, and received relevant ethics approvals (see Supplemental Appendix E).

First, our sample includes threads discussing the genocide of Jewish people during World War II, known as the Holocaust. This long-standing contentious topic has attracted many conspiracies developed by online contrarian communities (Bright, 2022), specifically in '*r/conspiracy*'. We collected data from all the threads in this community with higher engagement rates (defined as more than 200 comments) and containing the word 'Holocaust' in a title. As two identical threads discussed David Cole, one of the key voices in the contrarian Holocaust community (see Appendix A in Supplemental Material), we focused on the most commented thread. This strategy allowed us to sample two threads focusing exclusively on the Holocaust, in addition to three threads discussing the Holocaust and the next topic of interest, the Holodomor.

Second, the Holodomor was orchestrated by the leadership of the Soviet Union in order to crush peasant resistance to the collectivisation and maximise grain delivery for Soviet industrialisation, as well as to eliminate the Ukrainian independence movement (Applebaum, 2017). The official Russian propaganda denies that this constituted genocide, while more than 30 countries and the EU have officially recognised the Holodomor as genocide. We again selected all the threads in the *r/conspiracy* community that contained 'Holodomor' in the title and that had more than nine comments. This strategy resulted in four threads focusing on the Holodomor.

Finally, we selected a discussion, 'Are we witnessing genocide?', that focuses on the COVID-19 vaccination policy as an act of genocide. Perhaps this example is too recent to be regarded as history; however, it offers an instructive complement to the two other cases about the dynamic of knowledge, and the pandemic will, of course, be discussed as history in the future (Baldwin, 2021).

Online Communities: Data Analysis

Our analysis of online data builds on existing approaches to study contrarian expression on Reddit, specifically around climate change (Oswald & Bright, 2022). As explained earlier, we collected data on three history-related cases using Reddit's official application programming interface in November 2022. This data comprised 2014 user comments (our unit of analysis) on three genocide-related themes from 10 threads.

The analysis consisted of three stages. First, we took a random sample of 1725 comments from the collected data and coded them as either featuring 'contrarian statements' or

Table 1. Descriptive Statistics of the Dataset Analysed.

	Manually coded
Contrarian statements (mean)	0.09
Contrarian statements (SD)	0.22
Distortion (mean)	0.05
Distortion (SD)	0.29
Score (mean)	8.23
Score (SD)	41.03
Number of replies (mean)	1.08
Number of replies (SD)	1.48
Number of observations	1725

Notes. For the definition of contrarian statements and distortion variables, please see Supplemental Appendix B. Score is the sum of up- and down-votes received from users.

'distortion' (Supplemental Appendix B). This simplified approach cut across many nuances of historical debates: many contributors expressed a mix of scepticism about the international consensus in relation to one or another topic while supporting other parts of the consensus. Nevertheless, this simplification was essential for our study; it captured key divisions within communities. The manual coding allowed us to address the peculiarities of natural language that algorithmic solutions are often unable to capture effectively, such as humour or images.

Second, we conducted a pilot coding exercise of approximately 10% of our random sample in three steps, which allowed a team of four postgraduate research assistants and one of the authors to compare their coding experiences before adopting the final coding template. Following this, the full random sample was coded. The coders compared their article coding and resolved any discrepancies through conversation with the broader research team. See Figure C1 in Supplemental Appendix C for histograms of the results of this analysis.

Third, we analysed our results descriptively (Table 1) and conducted data diagnostics (see Supplemental Appendix D). This analysis entailed a need to log-transform the outcome variables for our models examining the reinforcement of contrarian expression via up-voting (M1) because of the skewed distributions observed. We built these models using a linear regression with multiple group fixed effects. We included fixed effects on the time when a comment was posted to control for the temporal effects of the data, as posting varied over time. We also included fixed effects on the theme of the subreddit where a comment appeared.

For M2 models that tested how users were replying to contrarian expression, we used a generalised linear model with a negative binomial distribution due to the presence of overdispersion in the data, controlled by the time of the day and the topic of a subreddit where the comment appeared. This approach allowed reducing overdispersion and address multicollinearity identified among the dummy variables for the day of the week and year. For M3, we utilise a PWPGap

Time model (Amorim & Cai, 2015) to analyse the relationship between engagement and the likelihood of commenting again. The model considers complete comment histories, clusters standard errors at the author level and includes a number of replies the comment received as a control variable. We used HC1 robust standard errors in all models. See Supplemental Appendix D for details.

There are three key limitations to the methodology detailed so far. First, Reddit is a US-centric space. Nevertheless, it is one of the most visited social media platforms globally, and our impression of the data suggested it was not particularly US-centric. Second, the scope of subreddits and topics was limited (we could have examined more history topics, for example), and any quantitative analysis of texts is bound to lose some of the nuances of the discussions. Therefore, qualitative analysis of the Reddit data or interviews with Reddit participants would have benefited our analysis. Third, as Supplemental Appendix C discusses, Reddit data is rarely complete over time, as with any mainstream social media.

Expert Interviews

To contextualise our findings on online communication, we solicited the views of experts in the public understanding of history, online communication, and digital media through semi-structured interviews. Our research questions focus specifically on the dynamics of online contrarian communities, so the primary inclusion criterion was demonstrated expertise in how historical narratives and public knowledge are shaped and contested in digital environments. Quantitative analysis on its own would be rather uninformative without giving a sense of how experts make sense of contrarian expression about history online, specifically in relation to the topic of genocide. Furthermore, without contextualization of how experts think about how the topic is discussed online, also compared to other topics and to how it is discussed in other media, our analysis would lack a sense of why such online discussions are significant for how the public engages with history. We interviewed 10 experts, whom we selected based on their direct knowledge of the subject, and used the snowball approach to obtain a wide range of views. See Supplemental Appendix E for a list of the interviewees and the interview protocol.

Results

Contrarian Expression

This section presents our analysis of how users in Reddit contrarian communities typically react to contrarian expression about genocide. Descriptive statistics of our dataset are presented in Table 1. We found that 10% of the coded content tended to contain distortion statements, while contrarian comments constituted 6% of the coded dataset.

In other words, more comments tried to direct a conversation into questioning the accepted historical interpretation in a subtle way rather than outright opposing this interpretation. Hence, contrarian expression is only a minority in these history-debating communities. We also see that the conversational content we analysed was less engaging than other types of contrarian expression: on average, every comment we studied received only one reply and eight scores. In Oswald and Bright's (2022) study of climate scepticism on Reddit, by way of comparison, each comment received seven replies and 14 scores on average.

Before presenting further quantitative results of our study, we provide a few full examples of data points in a context that demonstrate the results based on our thematic analysis of the conversations. The examples are mostly (mostly, because some comments are not revealing) chosen because they are among content pieces in our dataset that received the most engagement as measured by the number of comments and scores.

The community comprising users who discuss the Holocaust sometimes refers to prominent voices as 'researchers' building up 'revisionist institutions'. These users want to sound like academic historians referring to 'historical' publications and arguments. One of the comments in the Holocaust thread shows this tendency:

For anybody wanting to read a good revisionist study of Treblinka by actual, sober, publishing researchers, there is Treblinka. Extermination Camp or Transit Camp? [link trimmed by Authors] (0 comments, 7 scores)

In contrast, those who support the existing scientific consensus around the Holocaust are often described in this community as 'anti-revisionist':

The anti-revisionist laws in many countries are foul and stupid. They exist as part of a larger picture of political correctness that also outlaws 'Islamophobia', 'homophobia', 'racism', etc. Thankfully, here in the U.S. the First Amendment has held firm in the face of attempts to limit political speech. Other nations are not so fortunate. (2 comments, 9 scores)

Perhaps the major claim discussed around the Holocaust concerns the number of victims, rather than denying the fact of the Holocaust outright:

The Korherr Report (Himmler's statistician) puts the number of murdered Jews at a little under 2.5 million by early 1943. Deniers can't account for those Jews. It's because they were indeed dead. Add to Korherr's figure the Jews who died after the extermination camps were closed in '43, and you'll get a ROUGH figure of 3 million. It's a Holocaust, no doubt. But not 6 million. (3 comments, 3 user scores)

However, the community is not necessarily dominated by Holocaust revisionists. The following is the highest-ranked

comment about the Holocaust, which expresses no contrarian views:

Welp, I've read all that I need to read. This discussion is over. I really thought we had a live one here, not another disinfo Zioshill. (3 comments, 42 scores)

The Holocaust also emerges as a topic of other analysed themes on r/conspiracy, concerning vaccination as an act of 'genocide'. This example is the most commented comment identified as opposing the accepted scientific consensus in our dataset:

If we get this vaccine every 4-6 months, what makes you guys think they won't put in graphene/nanobots and take out whatever race or whatever their agenda is slowly? Yes this is a genocide, of who.. it is unclear. But Klaus Schwab's family was friends with Hitler. It is a very slow boil genocide. They created the virus, and they created the solution which is mandatory now to keep your job (8 comments, 11 scores)

The main theme that united many users expressing an opinion about the Holocaust and vaccination was a mistrust of the government or other institutions with authority. Yet the engagement with the topic was quite different between the two types of threads: the Holocaust content was much more rooted in a quite detailed discussion of history, involving exact numbers, dates, and places, while recent events in the news drove vaccination genocide debates.

The most engaged content in our dataset covered the questions of the Holodomor as an act of genocide. For example, this highest-scored comment also discussed the Holodomor and was identified as using distortion:

Don't forget as many westerners seem to've though that Winston Churchill took similar actions that caused millions to starve in Bengali India. It's not just a problem of Stalinism or communism, it's a problem of any people with excessive authority who believe their ethos to be utilitarian. [. . .] (3 comments, 56 scores)

The lowest-scored comment that we identified using distortion is presented below. It compared the Holodomor to the Holocaust:

[. . .] there is a difference between making big mistakes while trying to industrialise a backwards feudal society, and intentionally creating death camps for jews. (3 comments, -20 scores)

The highest-scored comment overall was also about the Holodomor. Just like most of the content in our dataset, it expressed no opposing or distortion views. However, the score it received was really unusual for our dataset (Table 1):

Its seriously disgusting the amount of people that refuse to admit this was intentional, when Stalin orders grain collected from

ukrainian peasants, when there was already a very poor harvest and orders it sent to the army and sold abroad that is fucking intentional. We as an entire socieity need to stop this rewriting of the history of the stalinist USSR. [. . .] (19 comments, 1,293 scores)

Against the background of these examples, we return to the quantitative analysis. Table 2 presents fixed-effects linear regression models, with each observation representing a comment. An additional control is comments without a URL. The score outcome variables are log-transformed, with exponentiated coefficients representing percentage changes.

Hypothesis Testing

To investigate whether contrarian expression is reinforced via up-voting while other types of expression are devalued (H1), we examined the relationship between comment content and comment score. The results indicated that comments containing contrarian statements ($b=0.04$, $p=.388$) or distorted statements ($b=0.02$, $p=.491$) do not receive significantly higher scores than other types of comments across the analysed discussions. In addition, this model has several limitations, as discussed in Supplemental Appendix D. These results provide insufficient evidence to support H1, suggesting that other types of expression are not necessarily devalued.

H2 posits that content featuring contrarian expression receives more replies than non-contrarian content. Comments with higher distortion levels ($b=0.19$, $p=.02$) or containing contrarian statements ($b=0.22$, $p=.037$) received significantly more replies. While this quantitative evidence is modest and should be interpreted with caution, it indicates that contrarian content may trigger more active discussion, supporting H2.

H3 posits that individuals engaging with contrarian expression are more likely to continue their engagement with the community. We measure this engagement through commenting activity. The analysis of variation in the effect of engagement with contrarian comments on a chance of posting again over time (M3) revealed a statistically significant association between engagement with contrarian statements and a shorter time until a user's subsequent comment (HR=1.36, $p=.042$). This suggests that users who engage with contrarian statements tend to post again more quickly. However, no statistically significant relationship was observed for engagement with comments of a milder expression type, distortion (HR=0.87, $p=.23$).

A proportional hazards test revealed that this effect varies over time (Figure 1), showing an increase in likelihood to comment again within the first 10min before stabilising around the mean. This suggests that exposure to contrarian statements, though not distortion, is associated with increased continued engagement, providing partial support for H3.

Table 2. Reactions to Contrarian Expression.

	Reinforcement (log) (M1)		Replies (M2)		Comment Again (log) (M3)	
Distortion	1.02 [0.90, 1.23]			0.19* [0.03, 0.36]	0.87 [0.42, 1.23]	
Contrarian statements		1.04 [0.89, 1.34]	0.22* [0.01, 0.43]			1.39* [1.02, 4.00]
Holodomor			0.20* [0.05, 0.36]	0.22** [0.06, 0.38]		
Vaccination			0.25** [0.09, 0.41]	0.26** [0.10, 0.42]		
Num. Obs.	1647	1647	1596	1597	768	768
R2	0.490	0.490	0.118	0.117		
R2 Adj.	0.168	0.168	0.098	0.097		
RMSE	0.24	0.24	0.92	0.92	0.84	0.83
Log. Lik.			-1946.2	-1950.2		
F			3.270	3.294		

Notes. 95% confidence intervals are reported in square brackets. For the definition of contrarian statements and distortion variables, please see Supplemental Appendix B. Reinforcement refers to the measurement of the balance of up- and down-votes a comment received. We calculated the probability to comment again for users who engaged with comments containing contrarian statements.

* $p < .05$, ** $p < .01$.

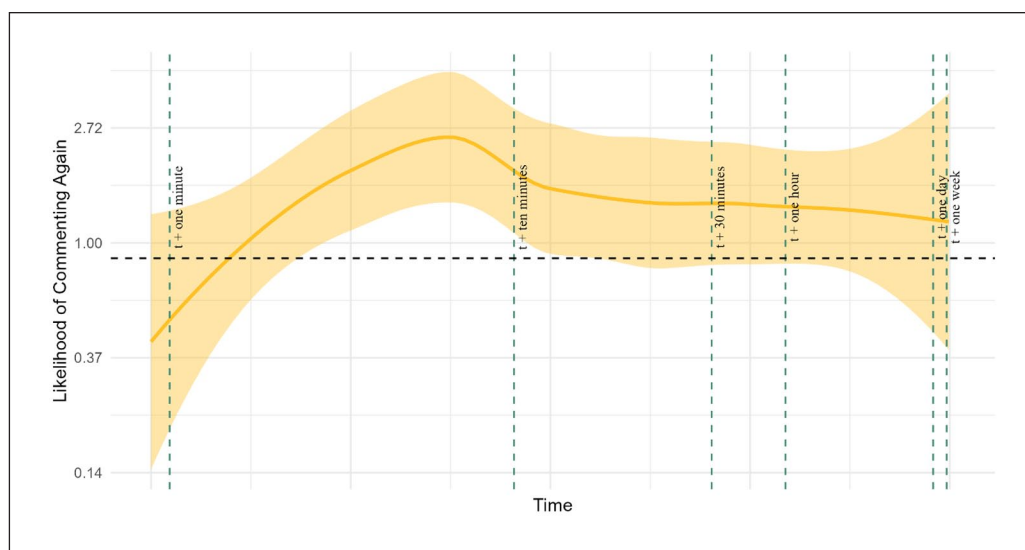


Figure 1. Probability to comment again.

Notes. Visualising users who engaged with comments containing contrarian statements only (M3, contrarian statements). The dashed horizontal line indicates the average effect of a probability to comment again; the orange area indicates the 95% confidence interval.

Expert Interviews

The discussion below focuses on three themes that emerged through thematic analysis of our interviews based on grounded theory: contrarian expression around history topics online, the role of platforms in spreading contrarian expression and possible solutions that can mitigate this process.

Contrarian Expression and History Online

The interviewees most often discussed contrarian expression about historical topics in the context of nationalism,

identities, populism or education. The relevance of historical narrative for politics was frequently emphasised, as the experts believed that it is hard to have a political argument without a historical narrative (Faisal Devji), ‘while particular kinds of conspiracy seem to appeal to particular political persuasions’ (Jane Winters). It was evident from the interviews that contrarian expression is widely spread in different contexts, both in the Global South and the Global North (Amogh Dhar Sharma), and used for political manipulation, though it is also the case that ‘many people regard history as a form of entertainment’ (Peter Furtado), which the experts linked to

the affective dimensions of media use, the delegitimization of expertise, and political populism:

Most historical conspiracy theories have some kind of political or social significance. And what they do first of all, is to challenge orthodox expert knowledge. They form a part of a more general phenomenon, which you can see in COVID-19 conspiracy theories, the American moon-shot in 1969, and many, many others. So it's part of a populist wave of challenging expert knowledge. (Richard Evans)

Nevertheless, experts' perceptions of manipulation and controversies often revolved around a belief in the continuation of many key debates and arguments in society, where 'what becomes controversial changes from time to time [. . .] and it's not clear how much of this is absolutely new and how much is returned' (Faisal Devji).

The experts did not tie the spread of contrarian expression to the introduction of the internet or social media exclusively. Instead, they saw the internet in the context of the evolution of technologies that are used to spread controversies. People in previous periods largely relied on word of mouth, and later the printing press for this purpose (Richard Evans). Still, in recent years, some qualities of the internet made conspiratorial content more widespread. Among them was the availability of free, open and abandoned content (Tobias Blanke), the availability of affordances for coordination of action (Josh Cowls), as well as the speed with which information spreads on social media and the possible 'volume' that it can reach:

The spread of manipulation online is not technology driven, but I think it is significantly technology assisted in terms of the speed and scale and availability. (Jane Winters)

As a result, the standard of public discourse 'has declined drastically: [. . .] it has become less courteous, less informed, and less rational', the historian Richard Evans believes.

The topic of the Holocaust is a good example that reflects all these aspects. The experts frequently linked this topic to the questions of the use of identity and nationalism in local contexts, where the discussion of the Holocaust is 'tied to rising nationalism' in 'countries like Poland and Hungary' (Jane Winters), as well as lack of education and the remoteness of this event:

It's easy to believe in some way that it didn't happen, cause how could it have? It was so huge, surely it can't have been that terrible. [. . .] The further away you get from it, the easier it is for people to say, well, it was distorted, it was exaggerated. So some of it is definitely around that educational deficit and the living memory question. And it's an uncomfortable topic. (Jane Winters)

At the same time, the experts suggested that people who deny the Holocaust would not really call themselves

'Holocaust deniers', for at least three reasons. First, they try 'to legitimise themselves by using terms that appear to be mainstream', such as 'historic revisionism' (as we saw in the previous section, 'revisionism' was often discussed on Reddit) – though 'all historians are revisionist' to some extent (Richard Evans). Second, some countries, like Austria and Germany, have introduced legislation that regulates expression around the topic of the Holocaust, so Holocaust deniers 'moderate' their language for fear of 'breaking the law' (Josh Cowls, Peter Furtado) – though not all experts found regulation effective in addressing the problem. Finally, the language of 'historical revisionism' has also become more sophisticated. Jane Winters compared this language to how 'the right in the US deals with school shootings', when instead of denying a tragedy has happened, interest groups 'try to minimise' the consequences and possible countermeasures.

The Role of Platforms

Almost every interviewed expert, unprompted, referred to the concept of 'gatekeeping' when discussing the evolution of public communication about history:

The internet and social media help to bypass traditional gatekeepers of opinion information. [. . .] What we're facing now is a period where disinformation spreads through the internet equally with information. A subject that we haven't really mastered yet we haven't decided where freedom of speech should be restricted. (Richard Evans)

However, some experts argued that platforms already function as new gatekeepers because they 'had to become more editorial' (Niels Bruegger). As a result, they can 'direct people to controversial content [. . .] without you being aware that is happening' (Jane Winters). Interviewees' prevailing assessment of this 'new gatekeeping' seems negative. Josh Cowls, who has researched fringe online platforms, observed that key gatekeepers on many platforms are human moderators, and communities on platforms like Reddit where contrarian content is more likely to spread are 'probably moderated by people who have a passion for conspiracy theories and are much more willing to believe them. And so moderation isn't going to be very effective in that case'. At the same time, 'the standards of historic debate can be quite good and less algorithmically affected' on history-dedicated and heavily moderated communities on Reddit compared to other popular platforms (Ales Bystryk). This can be explained by the 'pseudonymity' of its users, as well as the less-intrusive algorithmic curation in its timelines (Josh Cowls).

The experts were not fixated only on a gatekeeping function when explaining the spread of contrarian expression in online communities like Reddit. First, they emphasised that online contrarian communities will always be around, as people like 'to belong to communities' with established

norms, including the norms of expression (Ian Milligan). Second, they observed that contrarian expression might seem to dominate or be prominent in some online spaces due to ‘clustering’ (Josh Cowl), whereby much comment is shared online independently rather than being somehow organised in a coordinated way.

Third, some experts observed that platforms are full of ‘sensational and clickbait’ content (Peter Furtado), with debates being reduced ‘to the shouting of slogans [. . .] because you cannot expand your views at length, [. . .] and whoever shouts loudest seems to be the most listened to’ (Richard Evans). In this environment, communicating contrarian expression may be more appealing and persuasive than fact-based information (Jane Winters). Finally, the experts believed that many members of these communities are ‘not interested in the dialogue, they just want to continue repeating what they have already decided’. At the same time, they do not ‘have any good reason to connect to other actors unless it was something that could support their view’ (Niels Bruegger).

Solutions

Most of the experts we interviewed advocated greater regulation of platform companies that would focus on content moderation (Ian Milligan), a different model for advertising as a means of generating income (Jane Winters) and real competition and choice on the platform market by ‘fragmenting large corporations’ (Faisal Devji). Many interviewees saw ‘no value in endless arguments on social media’ (Faisal Devji), especially

in the case of conspiracy theorists, [who refuse to] to modify or change, let alone abandon their own arguments in the face of the evidence. And that’s why you can’t really argue with them. [. . .] It is very important that that debate takes place within the limits of what is allowed by the evidence (Richard Evans).

The interviewees argued that solutions could not be simplistic and should consider a humanities perspective instead of the exclusively technological one that seems to dominate the ‘solutions debate’ (Tobias Blanke). For example, web-historian Niels Bruegger argued for the regulation that should focus on ‘information handlers’ rather than on the web in general, as opposed to a platform-only focus. Jane Winters suggested bringing people formerly involved in the spread of conspiracies to work with contrarian communities, just like

former gang members quite often take part in youth work to try [to] keep people away from gang violence because they can speak to them as someone who has been in their position and understands them.

It is hard to overstate how much the experts emphasised the importance of education, school curricula and information literacy in addressing unverified information and conspiracies.

Discussion and Conclusion

Our findings add to the understanding of the spread of misleading information in online contrarian communities (Krafft & Donovan, 2020; Lewandowsky et al., 2019; Oswald & Bright, 2022). We demonstrate that when controversial topics are debated in such communities, their users do not necessarily reinforce contrarian expression if it is injected into their debates (H1). Instead, other expression types can be more prominent and engaged across these spaces.

Hence, just like some other types of contrarian communities, history-focused threads on Reddit are hardly fully homogeneous ‘echo chambers’ (Bright et al., 2022; Dubois & Blank, 2018; Herasimenka et al., 2023). Instead, these are spaces with vivid debates about controversies, at least when it comes to a moderated online platform.

However, it may be that the shared stance of the community is not allegiance to a specific contrarian view, but rather a commitment to a form of radical scepticism. In this framework, the rejection of any established claim – whether from the mainstream or fellow contrarians – is the very activity that draws online communities together, but this would also be contentious. Moreover, in contrast to climate change sceptic communities on the same platform (Oswald & Bright, 2022), contrarian expression is often only a minority in spaces where debates about history controversies take place.

We also show that contrarian expression is associated with more debate in the form of comments (H2), and users previously engaged with contrarian statements may be prompted to engage with a community again within a rather short period (H3), at least if they encounter contrarian rather than distorted statements. Just like many citizens remain uninformed rather than misinformed about events, misleading information does not dominate online communities with contrarian interests.

Our failure to confirm H1 contradicts the findings of Oswald and Bright (2022), who used a similar method to study contrarian expression about climate change. They found that more engagement between opposing views leads to contrarians digging in rather than moderating their position. Perhaps this reflects the difference between the domains of misleading information: statements covering climate change may generate more fractious engagement than history-related statements. In addition, our scope is broader and involves multiple themes and debates. On the other hand, we confirm H2 and provide partial support for H3 in line with the findings of Oswald and Bright: contrarianism leads to more engagement through comments.

This mix of findings has direct implications for our normative aim, as it demonstrates the potential role of platform engagement mechanisms in mitigating misleading information in the history domain. The finding that contrarian expression is likely to generate more debate suggests these are not passive bubbles; instead, such views are being actively contested. Therefore, our research shows that more focused debates, discussions and replies potentially help rather than damage efforts to prevent inaccurate information dissemination, even in relatively lightly moderated contrarian communities, by ensuring contrarian views face scrutiny.

Our expert interviewees pointed to many of the dysfunctions in the information policies concerning online history and offered a mix of solutions. Solutions like content moderation or content removal are bound to involve trade-offs, and this is an area of active policymaking and research (Schroeder 2025). Reddit encourages the community self-moderation of highly engaged discussions, though how exactly moderation happens varies between communities. Still, Reddit is likely to remain a space where contrarian expression is present, just like in other sections of the public sphere. However, as long as this does not produce illegal content, platform affordances for moderation may make this a source to add to the informed diversity of public expression.

There is no scope here to relate our findings to the extensive debates about objectivity in history, but the nature of these debates can be highlighted by a book first published in the late 1990s by one of our expert interviewees, which crystallised the disagreements among historians, 'In Defense of History' (Evans, 2018). The context of the book is important since this was the time when 'postmodernist hyper-relativism', as Evans refers to it (2018, p. 254), perhaps reached its peak. These debates have receded since then, but they long pre-date the discussions in the 1990s, and are bound to continue. Evans' book gained extensive coverage (pp. 346–347). He notes that he was attacked for his 'defense' not just by the postmodern relativist left-wing among the historical profession, but also by conservative right-wing historians who thought that his version of defending historical objectivity had too much in common with those favouring progressive policies.

This controversy goes beyond the scope of this article, but we can highlight one passage which gets to the heart of how historical objectivity relates to discussions of genocide: 'I am arguing that historical relativism, the idea that the truth of history can never be discovered, makes it impossible to refute Holocaust deniers' (2018, p. 312) – or, we would add, genocide deniers or purveyors of ideas linking genocide and vaccination. Academic history is, of course, not the same as popular history writing or the political or educational uses of history, though the two overlap in works of popular history by academic historians. The relations between them can be put in an overly simplified way as follows: academic history

may be the redoubt – the repository of ground truth – which prevents the manipulation – the backdrop against which popular understanding of history takes place.

There are parallels here with climate change science, where climate scientists inform or provide the backdrop against which the public or popular understanding of climate change takes place, and where contrarian expression is also rife. Thus, truth (not in the sense of a final truth but one based on being open to evidence), based on expert knowledge, needs to be part of a diverse public sphere in which open engagement and discourse can take place and which is not distorted or skewed or dominated by certain or narrow partisan agendas. The non-gatekept internet enables different groups to try to dominate the discussion, but it also allows challengers to emerge who can criticise dominant agendas. These tussles between groups are needed for societies to thrive and improve.

How different countries have dealt with their public understanding of history can also be documented historically, as with Germany and Austria and their very different approaches to their Nazi and Holocaust past (Art, 2005). Examples of countries whose governments have recently tried to make the public understanding of history conform more to narrow partisan agendas include China (Mitter, 2020) and India (Sarkar, 2019). Popular historians have provided accounts of how the history and identity of nations are interrelated (Furtado, 2017), and multiple groups try to engage the public with history.

But a tension remains which we cannot resolve here, between the truth acceptable to the careful judgement and evidence-based analyses of historians, the need for public understanding of history which should be informed by historians, the manipulative politicisation of history which must be guarded against and legitimate contention in relation to history which now also takes place online and which should take place against the background of vigorous and open debate informed by historians' insights. Furthermore, as we have mentioned, the need for the strictest content moderation is currently the subject of much debate (Gorwa, 2024; Schroeder 2025). Our argument is not that such contrarian expression should be immune from moderation; rather, our findings on engagement suggest that fostering counter-speech can be a powerful tool alongside it. We have pointed to some of the dysfunctions of online expression in the case of history which will be subject to how these debates are resolved more widely.


Historians will continually revisit discussions of genocide. Trentmann's recent history of Germany puts it as follows: 'Some writers claimed that the singularity of the Holocaust was a clever ploy by the German, American, and Israeli 'elites' to obscure the West's own racist violence, by silencing the views of other genocides. As history, this is plain wrong' (2023, p. 668; referencing the historian Moses (2021)). There will continue to be different approaches to

history and models of online historical discussion and knowledge. The internet can be a tool for publics to engage with history, with social media functioning as gatekeepers in different ways. We have shown how discussions about history take place online, including what experts think about this phenomenon. Our analysis contributes to understanding broader debates about online knowledge. It also sheds light on how it may be possible to avoid the manipulation and distortion of historical events.

Acknowledgements

This research was assisted by Kuldeep Bhattacharjya, Amanda Curi, Dina Dong, Edward Jones and Alvin Lam. The authors gratefully acknowledge the support of Pixstory. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the University of Oxford or our funders.

ORCID iDs

Aliaksandr Herasimenka  <https://orcid.org/0000-0002-5876-5562>
Ralph Schroeder  <https://orcid.org/0000-0002-4229-1585>

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Pixstory.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability statement

The anonymised dataset of the analysed Reddit content is available at <https://osf.io/2kwg8/>

Supplemental material

Supplemental material for this article is available online.

Note

1. The concept of misleading information lacks a universally accepted definition within the academic community, and more research is needed into the various terminologies (Kapantai et al., 2021). Multiple terms are in use, including ‘disinformation’, which is often specifically used for the subset of untrustworthy information that is spread intentionally (Kapantai et al., 2021). We use ‘misleading information’ as an umbrella term referring to any information that turns out not to be supported by facts and can inform contrarian expression.

References

- Amaya, A., Bach, R., Keusch, F., & Kreuter, F. (2021). New data sources in social science research: Things to know before working with Reddit data. *Social Science Computer Review*, 39(5), 943–960.
- Amorim, L. D. A. F., & Cai, J. (2015). Modelling recurrent events: A tutorial for analysis in epidemiology. *International Journal of Epidemiology*, 44(1), 324–333.
- Applebaum, A. (2017). *Red famine: Stalin's war on Ukraine* (1st ed.). Allen Lane.
- Art, D. (2005). *The politics of the Nazi past in Germany and Austria*. Cambridge University Press.
- Baldwin, P. (2021). *Fighting the first wave: Why the coronavirus was tackled so differently across the globe*. Cambridge University Press.
- Bright, J. (2022). *History under attack: Holocaust denial and distortion on social media*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000382159>
- Bright, J., Marchal, N., Ganesh, B., & Rudinac, S. (2022). How do individuals in a radical echo chamber react to opposing views? Evidence from a content analysis of stormfront. *Human Communication Research*, 48(1), 116–145.
- Case, D. O., Andrews, J. E., Johnson, J. D., & Allard, S. L. (2005). Avoiding versus seeking: The relationship of information seeking to avoidance, blunting, coping, dissonance, and related concepts. *Journal of the Medical Library Association*, 93(3), 353–362.
- Chinn, S., Hiaeshutter-Rice, D., & Chen, K. (2023). How science influencers polarize supportive and skeptical communities around politicized science: A cross-platform and over-time comparison. *Political Communication*, 41(4), 627–648.
- Dubois, E., & Blank, G. (2018). The echo chamber is overstated: The moderating effect of political interest and diverse media. *Information, Communication & Society*, 21(5), 729–745.
- Elgesem, D., Steskal, L., & Diakopoulos, N. (2015). Structure and content of the discourse on climate change in the blogosphere: The big picture. *Environmental Communication*, 9(2), 169–188.
- Evans, R. (2018). *In defence of history*. Granta Books.
- Farrell, H. (2016, May 19). The Chinese government fakes nearly 450 million social media comments a year. This is why. *Washington Post*. <https://www.washingtonpost.com/news/monkey-cage/wp/2016/05/19/the-chinese-government-fakes-nearly-450-million-social-media-comments-a-year-this-is-why/>
- Furtado, P. (Ed.). (2017). *Histories of nations: How their identities were forged*. Thames and Hudson.
- Gorwa, R. (2024). *The politics of platform regulation: How governments shape online content moderation*. Oxford University Press.
- Harmon-Jones, E., & Mills, J. (1999). An introduction to cognitive dissonance theory and an overview of current perspectives on the theory. In E. Harmon-Jones, & J. Mills (Eds.), *Cognitive dissonance: progress on a pivotal theory in social psychology* (pp. 3–21). Washington, DC, USA: American Psychological Association.
- Herasimenka, A. (2022). Movement leadership and messaging platforms in preemptive repressive settings: Telegram and the Navalny movement in Russia. *Social Media + Society*, 8(3), 1–14.
- Herasimenka, A., Bright, J., Knuutila, A., & Howard, P. N. (2023). Misinformation and professional news on largely unmoderated platforms: The case of Telegram. *Journal of Information Technology & Politics*, 20(2), 1–15.

- International Panel on the Information Environment. (2023). *Strategies for improving the global information environment: Results from a systematic review and meta-analysis* (No. SP2023.1). <https://www.ipie.info/research/sfp2023-1>
- Jost, J. T., & Amodio, D. M. (2012). Political ideology as motivated social cognition: Behavioral and neuroscientific evidence. *Motivation and Emotion, 36*(1), 55–64.
- Jungherr, A., & Schroeder, R. (2021). Disinformation and the structural transformations of the public arena: Addressing the actual challenges to democracy. *Social Media + Society, 7*(1), 2056305121988928.
- Kapantai, E., Christopoulou, A., Berberidis, C., & Peristeras, V. (2021). A systematic literature review on disinformation: Toward a unified taxonomical framework. *New Media & Society, 23*(5), 1301–1326.
- Klein, O., Spears, R., & Reicher, S. (2007). Social identity performance: Extending the strategic side of SIDE. *Personality and Social Psychology Review, 11*(1), 28–45.
- Krafft, P. M., & Donovan, J. (2020). Disinformation by design: The use of evidence collages and platform filtering in a media manipulation campaign. *Political Communication, 37*(2), 194–214.
- Lewandowsky, S., Cook, J., Fay, N., & Gignac, G. E. (2019). Science by social media: Attitudes towards climate change are mediated by perceived social consensus. *Memory & Cognition, 47*(8), 1445–1456.
- Linkevicius de Andrade, D., & Vasques Filho, D. (2022). Moderation and authority-building process: The dynamics of knowledge creation on history subreddits. *Internet Histories, 6*(4), 369–390.
- Marwick, A. (2018). Why do people share fake news? A socio-technical model of media effects. *Georgetown Law Technology Review, 474*(2).
- Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. Data and Society.
- Mitter, R. (2020). *China's good war: How World War II is shaping a new nationalism*. Belknap Press.
- Moses, A. D. (2021). *The problems of genocide: Permanent security and the language of transgression*. Cambridge University Press.
- Neff, G., Jordan, T., McVeigh-Schultz, J., & Gillespie, T. (2012). Affordances, technical agency, and the politics of technologies of cultural production. *Journal of Broadcasting & Electronic Media, 56*(2), 299–313.
- Oswald, L., & Bright, J. (2022). How do climate change skeptics engage with opposing views online? Evidence from a major climate change skeptic forum on Reddit. *Environmental Communication, 16*(6), 805–821.
- Robison, J. (2020). Does social disagreement attenuate partisan motivated reasoning? A test case concerning economic evaluations. *British Journal of Political Science, 50*(4), 1245–1261.
- Sarkar, T. (2019). How the Sangh Parivar writes and teaches history. In A. P. Chatterji, T. B. Hansen, & C. Jaffrelot (Eds.), *Majoritarian state*. Hurst Publishers.
- Schroeder, R. (2022). Can the Internet Advance the Social Good? Mapping the Landscape. <https://www.oii.ox.ac.uk/news-events/reports/can-the-internet-advance-the-social-good-mapping-the-landscape/>
- Schroeder, R. (2025). Content moderation and the digital transformations of gatekeeping. *Policy & Internet, 17*(2), e425.
- Sturm, T., & Albrecht, T. (2021). Constituent Covid-19 apocalypses: Contagious conspiracism, 5G, and viral vaccinations. *Anthropology & Medicine, 28*(1), 122–139.
- Trentmann, F. (2023). *Out of the darkness: The Germans, 1942-2022*. Penguin.
- Wojcieszak, M. (2010). 'Don't talk to me': Effects of ideologically homogeneous online groups and politically dissimilar offline ties on extremism. *New Media & Society, 12*(4), 637–655.

Author biographies

Aliaksandr Herasimenka is Lecturer in Data Science and Communication at University of Liverpool and Research Associate at Oxford internet Institute, University of Oxford, Oxford, United Kingdom. He studies the use of digital communication for political manipulation.

Ralph Schroeder was formerly Professor in the School of Technology Management and Economics at Chalmers University in Gothenburg (Sweden). He completed his PhD about Max Weber at the LSE in 1988. His publications include *Rethinking Science, Technology and Social Change* (Stanford University Press, 2007) and *Being There Together: Social Interaction in Virtual Environments* (Oxford University Press, 2010). He is also the author of 'An Age of Limits: Social Theory for the 21st Century' (Palgrave Macmillan 2013) and, with Eric T. Meyer, of 'Knowledge Machines: Digital Transformations of the Sciences and Humanities' (MIT Press 2015). He is currently Professor at Oxford internet Institute, University of Oxford, Oxford, United Kingdom.