Understanding Mass Influence

A case study of Facebook as a platform enabling mass influence
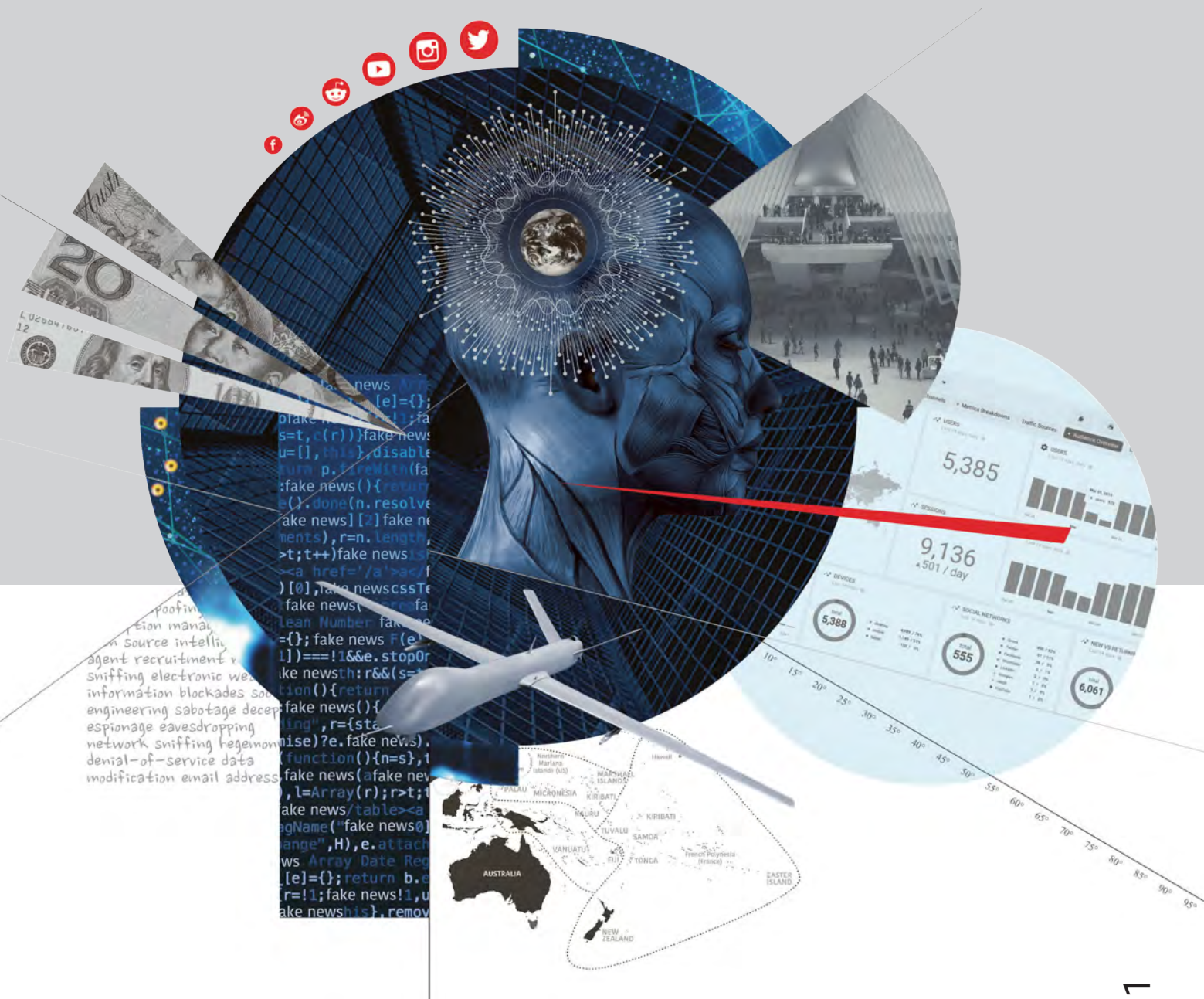
JULY 2021

Authors and Contributors

Authors

University of New South Wales
Dr Garry Young
Dr Peter Job

Contributors

Defence Science and Technology Group
Ms Laura Carter

Edith Cowan University
Dr Andrew Dowse
Dr Violetta Wilk
Mr Conor McLaughlin

Macquarie University
Professor Ben Schreer
Dr Brian Ballsun-Stanton
Dr Julian Droogan
Ms Lise Waldek
Mr Jade Hutchinson

The University of Melbourne
Professor Yoshihisa Kashima
Dr Jey Han Lau

University of New South Wales
Professor Monica Whitty
Associate Professor Stephen Doherty
Associate Professor Douglas Guilfoyle
Associate Professor Josh Keller
Associate Professor Rob Nicholls
Associate Professor Salih Ozdemir
Dr Raymond Wong

Contents

Figures

Introduction

TheFacebook – the company we now know as Facebook – was founded in February, 2004 by a group of Harvard University students, Eduardo Saverin, Andrew McCollum, Dustin Moskovitz, Chris Hughes and Mark Zuckerberg.[1] It initially served the Harvard student community but expanded in the years following to universities in the US, Canada and elsewhere. In September, 2006 Facebook opened to anyone aged 13 or over who wanted to join[2] and today is the world's most widely used social media platform. Facebook, Inc. launched on the US stock exchange in February, 2012.

On 31 December, 2020 Facebook claimed a global workforce of 58,604, up 30% on the previous year, with growth expected to continue.[3] In addition to the Facebook platform company assets include the photo and video sharing app Instagram, messaging apps Messenger and WhatsApp, digital gaming company Oculus VR, and others, including ad-serving, mobile applications, exercise apps, monetisation platforms and visual filtering gaming apps. Facebook also does business through a group of subsidiaries in a number of countries (see Appendix 1).

According to the 2020 Facebook Securities Exchange Commission (SEC) submission, worldwide there were on average 1.84 billion Facebook daily active users (DAUs) in December, 2020, an increase of 11% year-over-year. Monthly active users (MAUs) numbered 2.80 billion as of December 31, 2020, a 12% increase on 2019. Family daily active people (DAP) (i.e., individuals who visited at least one of the following – Facebook, Instagram, Messenger or WhatsApp – daily) numbered 2.60 billion on average in December, 2020, up 15% year-over-year. The figure for family monthly active people (MAP) was 3.30 billion as of December 31, 2020, up 14% year-over-year.[4]

Social Media News[5] reports that in January, 2020 Facebook had 16 million monthly active Australian users. This does not represent 16 million discrete users, however, as some people may have multiple accounts and other accounts belong to organisations. provides the following breakdown of Australian users by age (see Figure 1).

As a social media business, Facebook is a multisided platform. It has, however, a primary source of revenue: advertising. Facebook is a *for-profit* corporation (see Figure 2) working to maximise its shareholders' welfare. It does this by paying dividends to its shareholders or increasing its stock value, or both. Shareholder welfare is maximised when advertising revenue is maximised.

---

[1] Nicholas Carlson, "At Last – the Full Story of How Facebook Was Founded," *Tech Insider*, 5 March 2010. https://www.businessinsider.com.au/how-facebook-was-founded-2010-3?r=US&IR=T#we-can-talk-about-that-after-i-get-all-the-basic-functionality-up-tomorrow-night-1.
[2] "Open Facebook," *Forbes*, September 11, 2006. https://www.forbes.com/2006/09/11/facebook-opens-up-cx_rr_0911facebook.html?sh=77b79defa23c.
[3] "Form 10-K. Annual report pursuant to section 13 or 15(D) of the Securities Exchange Act of 1934. For the fiscal year ended December 31, 2020. Facebook, Inc.," (Washington: United States Securities and Exchange Commission, 2020), 11. https://www.sec.gov/Archives/edgar/data/1326801/000132680116000043/fb-12312015x10k.htm.
[4] Form 10-K, Facebook, Inc., 52.
[5] Social Media News. *Social Media Statistics Australia*, January 2020. https://www.socialmedianews.com.au/social-media-statistics-australia-january-2020/.
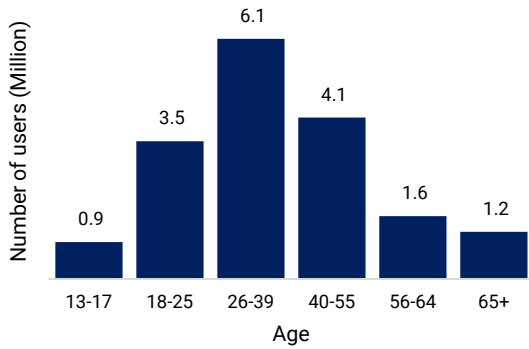
Figure 1: Number of Australian Facebook users by age
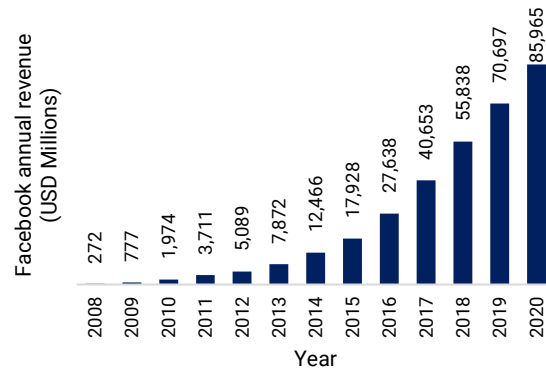
Figure 2: Facebook annual revenue

Below are three timelines delineating Facebook's position on key issues. The first provides examples of Facebook's aggressive acquisitions and suppression strategy, designed to help it monopolise the online social communications market. Recent opposition to this is also included. In addition, Figure 3 shows Facebook's position on user privacy which we see shift after the Cambridge Analytica story in 2018. The second timeline maps Facebook's changing attitude towards (allegedly) controversial content: from an initial reluctance to police content to a (more public) realisation and acceptance that social media platforms, including Facebook, need to be more accountable when it comes to enforcing standards. The final timeline illustrates Facebook's shifting position on whether it is a platform or a publisher.

Facebook has often resisted describing itself as a publisher. This is because, in the US, publishers are liable for the content they publish. Section 230 of the 1996 US *Communications Decency Act* provides Facebook with some protection, however, insofar as it allows a platform to make editorial decisions about content – to act like a publisher – without having to accept liability for all content published on its platform. Legislation has been introduced recently in the US to limit the application of Section 230. It cannot be used to absolve platforms of liability in cases of content related to sex trafficking, for example. There has also been talk of the current White House administration seeking to limit further, if not altogether abolish, Section 230.

17 Mar 2018
Cambridge Analytica scandal breaks

May 2018
Facebook responds to EU General Data Protection Regulation (GDPR) by moving users in Africa, Asia, Australia and Latin America from EU jurisdiction to the more lenient jurisdiction of US privacy laws.

Nov 2011
Facebook settles with US Federal Trade Commission regarding charges it violated users' privacy by allowing data to be shared and made public.

Aug 2017
The Global Internet Forum to Counter Terrorism (GIFCT), founded by Facebook, Google, Microsoft and Twitter. Aims to foster cooperation.

Mar 2019
Mark Zuckerberg claims to be "pivoting to privacy". Critics argue that this move is part of Facebook's strategy of owning the space of personal communication.

Sep 2006
Facebook becomes available to anyone, age 13+.

Mar 2014
Facebook purchases *Oculus VR* for US$ 2 billion.

Feb 2004
TheFacebook founded.

2004 — 2006 — 2011 — 2012 — 2013 — 2014 — 2015 — 2016 — 2017 — 2018 — 2019 — 2020

2006 - ongoing
Facebook employs a business model of aggressively acquiring or suppressing rivals.

Apr 2012
Facebook purchases *Instagram* for US$ 1 billion.

Feb 2014
Facebook purchases *WhatsApp* for US$ 19 billion.

Apr 2018
US Joint House Congressional hearing on privacy. Zuckerberg admits that Facebook has not done enough to prevent its misuse by malign actors. Promises to undertake reforms.

Feb 2019
The UK *Disinformation and 'Fake News'* report criticises Facebook for using its monopoly position to deny competitors access to data, causing their businesses to fail.

Dec 2020
The FTC, and 46 States and Districts across the US, accuse Facebook of maintaining a social networking monopoly and of anti-competitive conduct. Could result in the forced divestiture of Instagram, WhatsApp, and other assets.

Figure 3: Monopoly and privacy timeline

Feb 2008

Facebook criticised for hosting groups dedicated to anorexia. Facebook refuses to remove them simply because of their controversial nature.

Oct 2013

After deleting pages containing violent content (in response to criticism), Facebook reinstates them, stating that it would continue to host material which condemned not glorified violence.

Oct 2016

Facebook refuses to remove statements by Donald Trump on Muslim immigration despite accusations that they violate Facebook's rule on hate speech, arguing Facebook should not arbitrate truth for politicians.

Oct 2017

In relation to the 2016 US Presidential election, Facebook CEO, Sheryl Sandberg, states that if Russian activists were people not bots, Facebook would remove their content (in order to uphold freedom of expression). Facebook is a tech company not a news organisation.

May 2019

Facebook announces that users who break Facebook's community guidelines will face bans.

Oct 2019

Zuckerberg maintains Facebook is a tech company and should not police free speech.

2004　2008　2011　2012　2013　2014　2015　2016　2017　2018　2019　2020

Oct 2011

A Canada-based group – *Rape is no joke* – organise a petition demanding the removal of pages glorifying rape. Facebook refuses but updates its policy on hate speech.

Oct 2015

Facebook announces that it will be adding warning labels to some violent content.

Aug 2017

The Global Internet Forum to Counter Terrorism (GIFCT) is founded by Facebook, Google, Microsoft and Twitter to cover the spread of extremist and violent content.

Mar 2019

Christchurch Call to Action. Facebook and other tech companies sign on to the initiative to coordinate industry efforts to combat violence and extremism online. Zuckerberg agrees internet companies should be more accountable for enforcing standards.
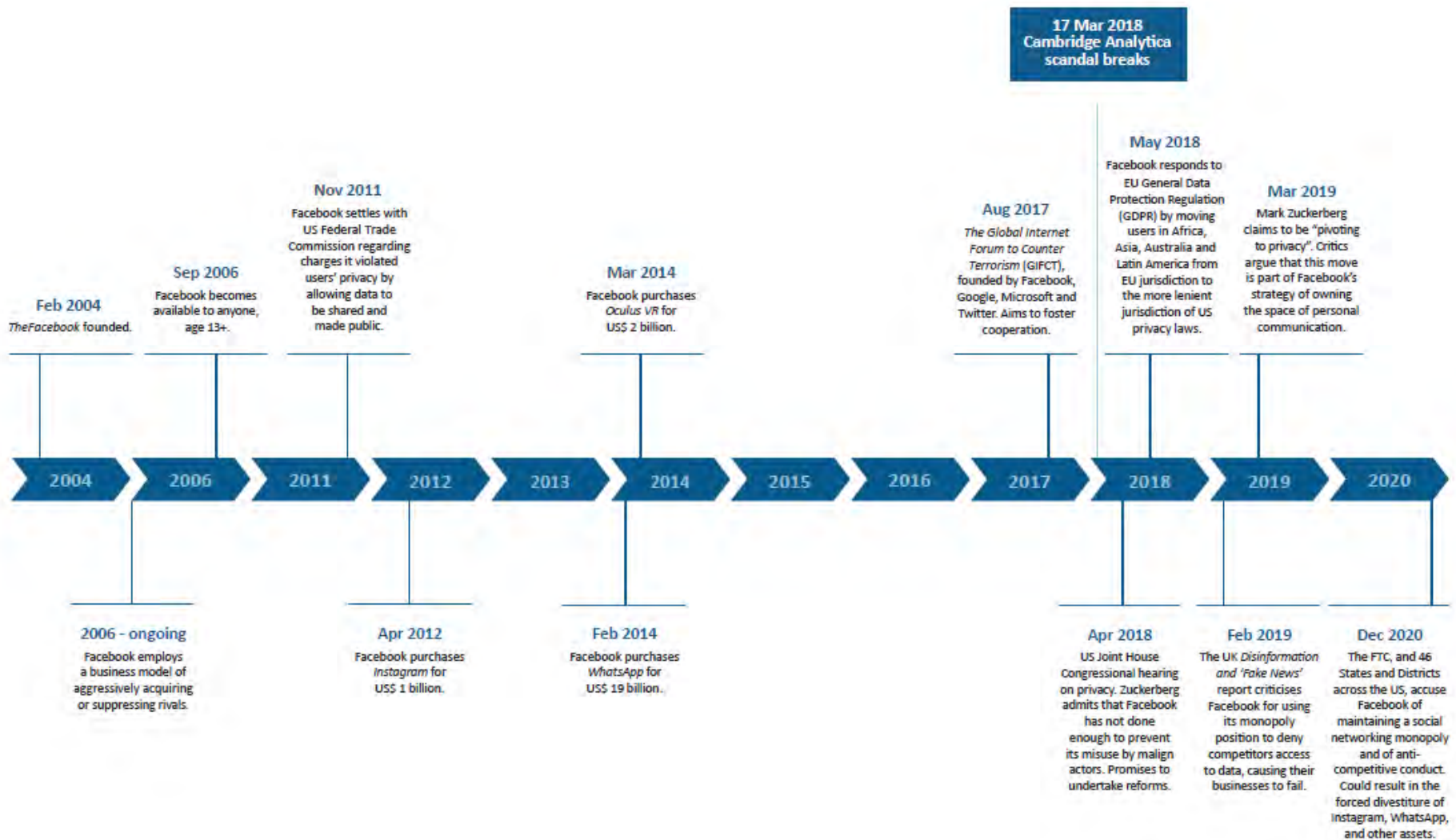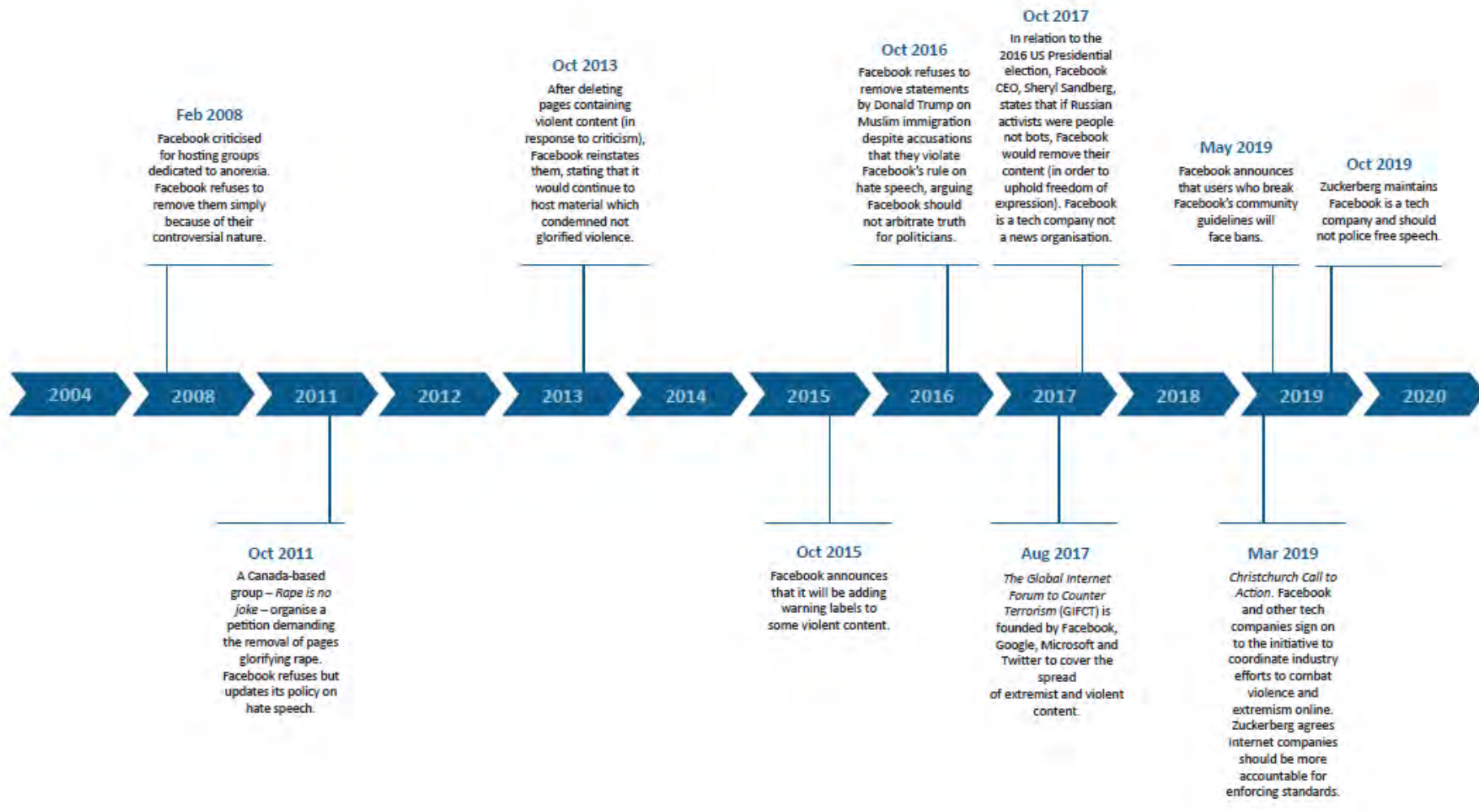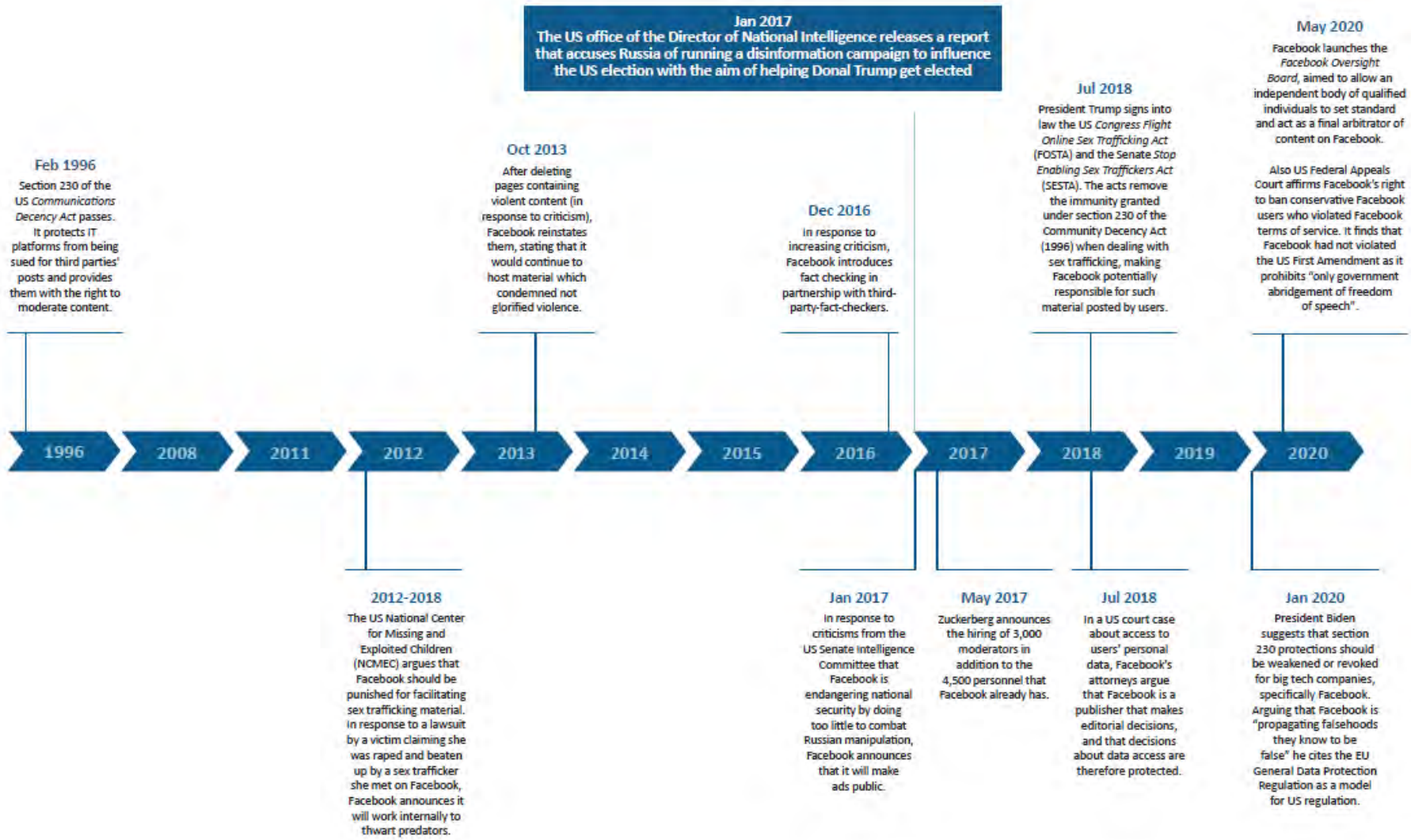
Figure 4: 'Controversial' content timeline

Figure 5: Facebook as 'platform or publisher' and mis-/disinformation timeline

Theme 1: Governance and Ethics

Research Question 1: What is Facebook's business model for operations, including its operating concept, financing arrangements, governance, and legal and ethical framework?

Facebook's operating concept

Facebook generates income primarily through selling advertising on its various platforms, including the Facebook platform, Instagram, Messenger and third-party affiliated websites on mobile applications. It currently dominates the digital advertising market with more than nine million advertisers and 19.7% of the digital advertising market,[6] with the potential to reach 32% of the world's population.[7] Facebook collects data from users and employs algorithms to create profiles and allow targeted advertising. Advertisers configure their ad campaigns through Facebook Ads Manager, which allows advertisers to define a target audience. Configuration parameters include location, demographic parameters (gender, age, etc), behaviours (device) and interests. Interest parameters are the most relevant to most advertisers, with hundreds of thousands of possibilities. Facebook earned 98% of its revenue from advertising in 2020, and 2% from other sources. Other sources include delivery of consumer hardware devices, fees from developers using its payments infrastructure and more.[8] [9]

The Facebook model has proved very attractive to advertisers, based on targeted access to Facebook's massive consumer database. Users may be targeted specifically for their likes and interests, making more efficient use of advertising resources. The more targeted the ads, the more Facebook can charge for them, creating a voracious appetite within Facebook for user data. Facebook has a strong incentive to keep users engaged – it allows more time for data collection and viewing ads. As Forbes analyst Len Sherman points out, it is a highly attractive business model, with no cost of goods sold, no marketing costs and no selling costs, creating a "trifecta of high scale *and* high growth *and* high profit margins unmatched by any high-tech company, including Google, Amazon, Apple, and Netflix."[10] It is, accordingly, a model that Facebook has no wish to abandon or modify in any significant manner.

Company structure and control

Facebook is incorporated in Delaware US, although its principal offices are in Menlo Park, California. Its stock is structured into two categories: Class A Common Stock, with 10 votes per share and Class B Common Stock with one vote. Holders of Class A Common Stock control the majority of voting power. On 31 December, 2020 there were 3471 stockholders of Facebook Class A Common Stock but as many are held by stockbrokers and other institutions, it is not possible to determine the number of individual stockholders. On the same date there were 32 holders of Facebook Class B Common Stock. Class A Common Stock has been listed on the NASDAQ global select market since 2012, while Class B Common Stock is not listed on any stock exchange.[11]

At the time of writing Facebook's board of directors consists of Chairman and Chief Executive Officer Mark Zuckerberg, Chief Operating Officer and Director Sheryl Sandberg, and seven additional board members. The two-tier stock structure concentrates control in the Board of Directors and CEO Mark Zuckerberg. As Facebook qualifies as a "controlled company" under corporate governance rules for NASDAQ listed companies, it is not required to have a majority of independent board members. Facebook has never declared or paid cash dividends. Its 2020

---

[6] Omnicore (2021) https://www.omnicoreagency.com/facebook-statistics/ (Accessed 12 February 2021).
[7] McLachlan, Stacy. "27 Facebook demographics to inform your strategy in 2021." https://blog.hootsuite.com/facebook-demographics/#:~:text=6.%20Facebook%20reaches%20the%20largest%20number%20of%20users%20aged%2013%2D17andtext=Instagram%20reach%20for%20the%2013,teens%20compared%20to%20last%20year. (Accessed 7 February 2021).
[8] Form 10-K, Facebook, Inc., 64.
[9] Angel Cuevas et al., "Does Facebook Use Sensitive Data for Advertising Purposes? Worldwide Analysis and GDPR Impact," *Communications of the ACM* 64. No 1 (2021) 64: https://arxiv.org/abs/1907.10672v1.
[10] Len Sherman, "Why Facebook Will Never Change Its Business Model," *Forbes*, 16 April 2018: https://www.forbes.com/sites/lensherman/2018/04/16/why-facebook-will-never-change-its-business-model/?sh=1520c38764a7.
[11] Form 10-K, Facebook, Inc., 49.

annual report to the US Securities and Exchange Commission (henceforth referred to as the 2020 Facebook SEC submission) stated that it did not intend to pay cash dividends in the foreseeable future and would instead retain future earnings to finance expansion of its operations.[12]

Financial trends

Since floating on the stock exchange in 2012 Facebook has expanded considerably in terms of its profitability, acquisitions, number of users and the role it plays in global society. Its Class A Common Stocks rose from US$42.05[13] to US$273.16 by the end of 2020.[14] That said, more recent Facebook stock performance has been mixed. The "privacy crisis" (discussed below) and other factors led to a dip in its share value in 2018, although recovery was appreciable from 2019. Facebook revenue grew by 21.6% in fiscal year 2020, a deceleration from the 26.6% rise during 2019. While the second quarter of 2020 was a low point for advertising growth, however, growth improved in the second half of the year.[15] Overall, in the five years to 31 December, 2020 stock performance was below the average for the NASDAQ Composite Index and the Dow Jones Internet Composite Index (DJINET), although above that of the S&P 500.[16]
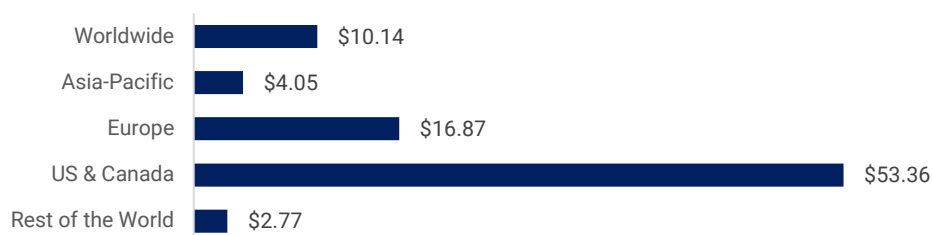

Figure 6: Annual revenue per user (US$) per region for the last quarter of 2020

According to the 2020 Facebook SEC submission, Facebook, Inc. revenue was US$85.97 billion in Fiscal Year 2020, up 22% year-over-year, and advertising revenue was US$84.17 billion, up 21%. Costs and expenses were US$53.29 billion. Income from operations was US$32.67 billion and operating margin was 38%. Net income was US$29.15 billion for a net profit margin of 33.9% in Fiscal Year 2020, with diluted earnings per share of US$10.09. As much as 45% or US$38.4 billion in revenue came from the US and Canada. The other 55% came from other regions across the globe. Capital expenditures, including principal payments on finance leases, were US$15.72 billion. Effective tax rate was 12.2%. Cash and cash equivalents and marketable securities were US$61.95 billion as of December 31, 2020. As said earlier, 98% of Facebook revenue came from advertising. Other revenue accounted for the remaining 2%, or US$1.8 billion.[17] Average revenue per user (ARPU) varies greatly by region (see Figure 6). Regions such as United States, Canada and Europe perform extremely well due to the size and maturity of their markets. Ad costs are higher than elsewhere due to Facebook's dominance in these areas, their greater consumer spending power and a consequent ability to charge more for ads. The following ARPU figures are for the last quarter of the 2020 US Financial Year (October-December):

— Worldwide: US$10.14

— US and Canada: US$53.36

— Europe: US$16.87

— Asia-Pacific: US$4.05

— Rest of World: US$2.7718

---

[12] Form 10-K, Facebook, Inc., 43-49.
[13] Michael J. de la Merced, "Facebook Closes at $38.23, Nearly Flat on Day," *New York Times*, 18 May 2012: https://dealbook.nytimes.com/2012/05/18/facebook-opens-at-42-05-in-debut-but-falls-quickly/?hp.
[14] Form 10-K, Facebook, Inc., 48.
[15] Form 10-K, Facebook, Inc., 19-20, 83.
[16] Form 10-K, Facebook, Inc., 49.
[17] Form 10-K, Facebook, Inc., 52, 96.
[18] Form 10-K, Facebook Inc, 56 (graph).

With revenue varying throughout the year, annual worldwide ARPU was US$32.03 in 2020, an increase of 10% on 2019. Growth varied by region. ARPU increased 18% in the US and Canada, 15% in Europe, 9% in the Asia–Pacific and was flat across rest of world. User growth (as opposed to ARPU) was more rapid in geographies with relatively lower ARPU and Facebook expects this trend to continue.[19] The majority of the company's revenue is generated by Facebook, Facebook Messenger and Instagram.[20] Facebook does not break down ad revenue by platform but reports indicate that more than a quarter of ad revenue comes from Instagram.[21]

Business model and acquisitions

Facebook employs a business model of aggressively acquiring or suppressing potential rivals. It purchased Instagram for US$1 billion in April, 2012. By 2019 more than 1 billion people were using the platform at least once a month and it has proved very profitable. Facebook acquired WhatsApp for US$19 billion in February, 2014. With 1.5 billion users in 180 countries it is now the most popular messaging service in the world. Oculus VR, an immersive virtual reality experience used in gaming, was acquired by Facebook for US$2.3 billion in cash and stock in 2014. Facebook has acquired a number of other companies over the years, as indicated in Appendix 2.[22]

The final report by the *Disinformation and 'Fake News'* enquiry by the Digital, Culture, Media and Sport Committee of the United Kingdom House of Commons (2019) criticises Facebook for using its monopoly position to deny competitors access to data, causing their businesses to fail. It cites the example of the Vine video app, launched by Twitter in 2013 and a potential competitor to Instagram Video. In an email to Zuckerberg and Facebook board members in January, 2013 a Facebook operative, Justin Okofsky, suggested that "unless anyone has any objections" he would shut down the app's access to Facebook data, which he then proceeded to do. In October, 2016 Vine announced that, partly due to the fact it could not grow its user base, it was discontinuing the app.[23]

The GDPR challenge

The European Union has attempted to enforce a more rigorous privacy regime on companies operating within its borders with the enactment in May, 2018 of the General Data Protection Regulation (GDPR). Companies may be fined up to €20 million, or 4% of global turnover (whichever is the higher), for the most serious violations of it. The law provides an array of consumer protections on Internet companies, including:

— The Right to obtain personal data. Users may obtain data the company has collected on them through a "data subject request".

— The Right to be forgotten. Companies must erase user data upon request.

— The Right to data portability. Users may transfer data to a new controller (such as another budgeting app).

— Lawful basis for processing. Consent to use data for one purpose does not imply consent for another.

— Affirmative consent. The controller is required to demonstrate that the data subject has consented to the processing of his or her personal data.

— Data protection by design. Systems must be designed to protect data.

— Profiling. EU users have the right to know how their personal information is being processed when an automated decision is made about them.

---

[19] Form 10-K, Facebook, Inc., 56-57.

[20] Gary Fox, "Facebook Business Model: How Does Facebook Make Money," *Gary Fox*, 8 March 2020: https://www.garyfox.co/facebook-business-model-makes-money/

[21] Sarah Frier, and Nico Grant, "Instagram Brings in More Than a Quarter of Facebook Sales," Bloomberg, February 5, 2020: https://www.bloomberg.com/news/articles/2020-02-04/instagram-generates-more-than-a-quarter-of-facebook-s-sales?srnd=premium.

[22] In November 2020, for example, Facebook bought *Kustomer* (a US loyalty and customer service platform) for US$1Billion, signalling to the market its intention to expand into customer service tools. This is a significant move as Facebook will be able to link its user profiles to the customer profiles, and will be able to charge businesses for customer service integration in this way. If this occurs, it will have huge ramifications on social commerce in the near future (Dr Violetta Wilk, personal correspondence, 21 May 2021).

[23] House of Commons Digital, Culture, Media and Sports Committee, *Disinformation and 'Fake News': Final Report, Eighth Report of Session 2017-19*, (London: United Kingdom House of Commons, 2019), 38. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf.

- Security requirements. Companies are required to provide the same level of protection for IP addresses and GPS information as they do for names and Social Security numbers.
- Parental consent. Users aged between 13 and 15 need consent from parents or guardians to share religious or political views. [24] [25]

Facebook has, in response, focused its efforts on pushing users to speed through consent processes, rather than reducing data collection. Constine describes the new processes as "a design that encourages rapidly hitting the 'Agree' button, a lack of granular controls, a laughably cheatable parental consent request for teens and an aesthetic overhaul of Download Your Information that doesn't make it any easier to switch social networks". He is particularly scathing of the parental consent process, noting that users only need to select one of their Facebook friends or provide an email address to obtain consent, making it easy to bypass. [26] Facebook has also made changes to move 1.5 billion non-EU users in Africa, Asia, Australia and Latin America, previously covered by European provisions, to the more lenient jurisdiction of US privacy laws. [27]

The challenge from Apple

Facebook is increasingly at odds with Apple over its use of user data, particularly in relation to Facebook products run on Apple platforms. Apple has been increasing its privacy protections, with CEO Tim Cook stating "we can no longer turn a blind eye" to the consequences of data collection and "a social dilemma cannot be allowed to become a social catastrophe". [28] In the first half of 2021, Apple will message Apple iPhone users asking whether they consent to apps on their platform, including Facebook, collecting their data. Apple will block data collection for those who decline. Facebook will attempt to pre-empt this by introducing a pop-up screen effectively urging users to opt-in to data collection, arguing it will allow Facebook to "continue to give people better experiences". Facebook has denied there is any trade-off between collecting data for the purpose of providing targeted ads and user privacy, claiming "in fact we can provide both". It says Apple's motive is to preference Apple's own targeted ads. With a considerable number of users accessing Facebook on iPhones, the real-world consequences will become apparent in the coming months and years.

Legal sanctions and challenges

As the 2020 Facebook SEC submission shows, Facebook is often the subject of regulatory investigation and legal procedures around the world. [29] The impact on Facebook's operations or profitability has, however, largely been limited. In October, 2018, for example, the UK's Information Commissioner's Office (ICO) imposed the then maximum penalty of £500,000 on Facebook for violating the UK's data protection law in relation to security issues around the harvesting of data. [30] The outcome of a number of recent cases in the US suggests that even substantially greater sanctions have only limited impact on Facebook's operations, or its incentive to make substantial changes regarding its use of data.

On 24 July, 2019 the SEC announced a settlement of charges against Facebook in relation to the Cambridge Analytica scandal, finding that Facebook had been aware from 2015 that Cambridge Analytica had misused Facebook user data (see Case Study 2) while, for more than two years, presenting such a scenario as only hypothetical in its public disclosures. Facebook paid US$100 million to settle the charges. [31] In a separate

[24] Kimberly A. Houser, and W. Gregory Voss, "GDPR: the end of Google and Facebook or a New Paradigm in Data Privacy?" *Richmond Journal of Law and Technology* 25, no. 1 (2018): 71-95, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3212210.

[25] Josh Constine, "A flaw-by-flaw guide to Facebook's new GDPR privacy changes," *Tech Crunch*, 18 April 2018: https://techcrunch.com/2018/04/17/facebook-gdpr-changes/.

[26] Constine, A flaw-by flaw guide to Facebook's new GDPR privacy changes.

[27] David Ingram, "Exclusive: Facebook to put 1.5 billion users out of reach of new EU privacy law," *Reuters*, April 19, 2018: https://www.reuters.com/article/us-facebook-privacy-eu-exclusive/exclusive-facebook-to-put-1-5-billion-users-out-of-reach-of-new-eu-privacy-law-idUSKBN1HQ00P.

[28] Taylor, Facebook v Apple: the looming showdown over data tracking and privacy.

[29] Form 10-K, Facebook, Inc., 36.

[30] House of Commons Digital, Culture, Media and Sports Committee, *Disinformation and 'Fake News': Final Report,* 20.

[31] "Facebook to Pay $100 Million for Misleading Investors about the Risks It Faced from Misuse of User Data," *U.S. Securities and Exchange Commission*, 25 July 2019: https://www.sec.gov/news/press-release/2019-140.

development the same day, the Federal Trade Commission (FTC) imposed a US$5 billion fine on Facebook for violating the 2011 FTC order by deceiving users about their ability to control the privacy of their data. The FTC also imposed new privacy restrictions on the company, establishing an independent privacy committee of Facebook's Board of Directors and removing unfettered control by Zuckerberg. Facebook must designate privacy compliance officers and conduct a privacy review of new or modified products. The settlement enhances the independent third-party assessors' ability to evaluate the Facebook privacy program and requires CEO Mark Zuckerberg to independently submit quarterly compliance certifications regarding the privacy program.[32]

The FTC voted 2-3 in favour of the settlement. While the FTC majority lauded the supposed "sweeping" extent of the new requirements, two Commissioners dissented. Commissioner Rohit Chopra argued that despite the US$5 billion penalty, the blanket immunity for Facebook executives and a lack of real restraints on Facebook business model "do not fix the core problems that led to these violations".[33] Commissioner Rebecca Kelly Slaughter argued that the civil penalty was insufficient given the injury to the public, the ability of the company to pay and the potential for the settlement to act as an effective deterrence.[34] With shareholders anticipating an outcome, the share price jumped after the company announced the two penalties. A combined US$5.1 billion sanction was apparently not sufficient for shareholders to lose faith in the company's ability to continue to operate its user data-based business model and produce profits for them in the future.

After a 16-month investigation into Apple, Amazon, Facebook and Alphabet, on 6 October, 2020, the Democratic majority of the US House Judiciary subcommittee on antitrust released a report concluding that the domination of these companies impacts on US democracy and the US economy. It suggested Congress implement changes to antitrust laws that could result in parts of the businesses being broken up.[35]

On 9 December, 2020 the FTC, in conjunction with a coalition of 48 states and districts across the US, filed parallel anti-trust lawsuits against Facebook, accusing the company of maintaining a social networking monopoly over many years by means of anti-competitive conduct. The FTC complaint claimed Facebook had systematically worked to eliminate threats to its monopoly, citing its 2012 acquisition of Instagram, its 2014 acquisition of WhatsApp and its alleged imposition of anti-competitive conditions on software developers.[36] At the time of writing the litigation is ongoing. According to the FTC, outcomes in federal court could include a number of options, including the forced divestiture of Instagram, WhatsApp and other assets. Other possibilities include prohibiting Facebook from imposing anti-competitive conditions on software developers, requiring it to seek prior approval for future mergers and acquisitions, or forcing it to allow users to post material simultaneously across competing social network platforms and permit friends lists to be exported to rivals, thereby increasing competition and making it easy for people to leave Facebook.[37]

Facebook has also been challenged by politically conservative elements in the US which accuse it of bias against conservative voices, and violation of free speech in its moderation practices and banning and suspension of users. Section 230 of the US Communications Decency Act protects websites and social networks from being sued for

[32] "FTC Imposes a $5 Billion Penalty and Sweeping New Privacy Restrictions on Facebook," *Federal Trade Commission, press releases*, 24 July 2019: https://www.ftc.gov/news-events/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions.
[33] "Dissenting Statement of Commissioner Rohit Chopra," *United States of America Federal Trade Commission*, 24 July 2019: https://www.ftc.gov/system/files/documents/public_statements/1536911/chopra_dissenting_statement_on_facebook_7-24-19.pdf.
[34] "Dissenting Statement of Commissioner Rebecca Kelly Slaughter," *United States of America Federal Trade Commission*, 24 July 2019: https://www.ftc.gov/system/files/documents/public_statements/1536918/182_3109_slaughter_statement_on_facebook_7-24-19.pdf.
[35] "Investigation of Competition in Digital Markets: Majority Staff Report and Recommendations," *Subcommittee on Anti-Trust, Commercial and Administrative Law of the Committee on the Judiciary*, (Washington: US House of Representatives, 2020): https://judiciary.house.gov/uploadedfiles/competition_in_digital_markets.pdf.
[36] "FTC Sues Facebook for Illegal Monopolisation," *Federal Trade Commission press releases*, 9 December 2020: https://www.ftc.gov/news-events/press-releases/2020/12/ftc-sues-facebook-illegal-monopolization.
[37] Sherman Bond, and Bobby Allen, "48 AGs, FTC Sue Facebook, Alleging Illegal Power Grabs to 'Neutralise' Rivals," *NPR*, 9 December 2020: https://www.npr.org/2020/12/09/944073889/48-attorneys-general-sue-facebook-alleging-illegal-power-grabs-to-neutralize-riv.

what third parties post but also provides them the right to moderate content and users on their platforms. Facebook maintains it has a right to remove hate speech and that it has always banned users who "promote or engage in violence and hate, regardless of ideology". The banning of a number of conservative Facebook users prompted outrage from some conservative US politicians,[38] and the conservative organisation Freedom Watch, along with YouTube personality Laura Loomer, brought a suit alleging Twitter and Facebook had violated the US First Amendment rights of users. However, on 20 May, 2020 a US Federal Appeals Court upheld a finding by a lower court that antitrust laws had not been violated and that the First Amendment did not apply to Facebook as it prohibited "only government abridgement of freedom of speech".[39]

The Australian media bargaining code and the Facebook Australian news ban

On 18 February, 2021 Australian Facebook users woke up to find that news and a variety of other material, including government health and emergency services, had disappeared from their Facebook News Feed. Facebook had removed the material in response to consideration in Australian Federal Parliament of a media bargaining code, officially called the Treasury Laws Amendment (News Media and Digital Platforms Mandatory Bargaining Code) Bill 2021, aimed at forcing Facebook to share revenues obtained from news articles posted on its platform by Australian media organisations with the organisations. Media platforms that did not reach agreements with news organisations may be "designated": that is, forced to negotiate through mediation and, ultimately, through arbitration.[40] The development made international news, much of it hostile to Facebook and sympathetic to Australia's position.[41]

On 23 February the Treasurer, Josh Frydenberg, announced that a compromise had been reached and that Facebook would reinstate Australian news pages. The new law would come into effect, with some of the provisions around "designating" watered-down. These included taking into account contributions a platform may have made to the sustainability of the Australian news industry through previous agreements and providing a longer period of notification regarding a decision to designate.[42] Commentators are divided about who came out on top in the skirmish.[43] Facebook's response to the media bargaining code demonstrates the lengths it is prepared to go to defend its business model. It judged that the damage done to its business in a relatively small jurisdiction such as Australia was worth the price of sending a message to potential reformers elsewhere.

Labelling and misuse of sensitive information

More broadly, Cuevas et al report that Facebook labels up to 67% of its users with potentially sensitive interests, sometimes at considerable risks to the users. In October, 2018, for example, Facebook labelled 540,000 users in Saudi Arabia with the ad preference "Homosexuality" and does this in other countries which have the death penalty for homosexuality. The researchers asserted that in the EU Facebook labels 73% of its users with potentially sensitive interests, possibly in contravention of EU law. These can include political opinions, sexual orientation,

[38] Jane Coaston, "The Facebook free speech battle, explained," *Vox*, 14 May 2019. https://www.vox.com/technology/2019/5/6/18528250/facebook-speech-conservatives-trump-platform-publisher
[39] Eric Larson, "Twitter, Facebook Win Appeal in Anticonservative Bias Suit," *Bloomberg*, 28 May 2020. https://www.bloomberg.com/news/articles/2020-05-27/twitter-facebook-win-appeal-over-alleged-anti-conservative-bias.
[40] "Treasury Laws Amendment (New Media and Digital Platforms Mandatory Bargaining Code) Bill 2021," Parliament of Australia, assented March 2 2021: https://www.aph.gov.au/Parliamentary_Business/Bills_Legislation/Bills_Search_Results/Result?bId=r6652.
[41] Queenie Wong, and Daniel Van Boom, "Facebook's News Ban in Australia: Everything You Should Know," *C/net*, 23 February 2021: https://www.cnet.com/news/facebooks-news-ban-in-australia-everything-you-should-know/.
[42] Amanda Meade, Josh Taylor, and Daniel Hurst, "Facebook reverses Australia news ban after government makes media code amendments," *The Guardian*, 23 February 2021: https://www.theguardian.com/media/2021/feb/23/facebook-reverses-australia-news-ban-after-government-makes-media-code-amendments.
[43] James Purtill, "Facebook thinks it won the battle of the media bargaining code - but so does the government," *ABC Science*, 28 February 2020: https://www.abc.net.au/news/science/2021-02-26/facebook-google-who-won-battle-news-media-bargaining-code/13193106.

personal health issues and other matters, potentially allowing malevolent actors to launch ad campaigns attacking specified groups of people based on it.[44]

Does Facebook's business model make its platform susceptible to manipulation by malign actors?

One of the most popular fake news stories trending on Facebook and Twitter during the 2016 US presidential election was the claim that the pope had endorsed Donald Trump's candidacy. On Facebook alone, this story received more than one million shares, and there were no doubt untold numbers who did not share but who were likewise exposed to the lie.[45] Thus, Marin states:

> Social networking sites such as Twitter or Facebook are very efficient channels for the propagation of misinformation because of the massive informational content shared by their users – content that they themselves did not author but only shared further. Regular social media users are responsible for most of the misinformation propagated on social networking sites… since misinformation would have much less harmful effect if it were not made visible by being shared by social media users. Without ignoring the effect that bots and for profit propaganda sites have in creating and sharing misinformation, the role of regular users in amplifying the storm of misinformation deserves further scrutiny because their well-intended acts of sharing content have an aggregated disastrous effect on the online information ecosphere.[46]

For Marin, the purported disastrous effect stems from two forms of epistemic harm: individuals may acquire false beliefs as a consequence of the mis- or disinformation they are exposed to and, more broadly, the ecosystem of online information is polluted with false content. Attention is also diverted away from genuine news items and stories of interest towards non-issues that then become issues (news) by virtue of the attention *they* receive, including being 'picked-up' by mainstream media. Yet Marin further notes that content shown to be false may elicit from those who share it the defensive response: "Sharing is not endorsing". While this is true, it is nevertheless indicative of the moral ambiguity surrounding the act of online sharing – particularly news items – with one's 'friends' and peers. This is because the epistemic threat posed by mis- or disinformation, in the case of newsworthy content, is deleterious to democratic processes that are best served by accurately informed citizens,[47] especially if, as some have found, "falsity travels with greater velocity than truth".[48] It can also have a negative impact on public health, or at least efforts to combat public health issues, as the COVID-19 pandemic and subsequent *infodemic* has illustrated.[49] Attempts are being made by organisations like WHO to combat COVID-19 infodemics,[50] as well as by social media platforms in relation to this and other (alleged) fake content. In 2017, for example,

[44] González Cabañas, Àngel Cuevas, Aritz Arrate, and Rubén Cuevas Communications of the ACM, 64, no. 1 (January 2021): 62-69.

[45] Prier, "Commanding the trend," 50-85.

[46] Lavinia Marin, "Sharing (mis)information on social networking sites. An exploration of the norms for distributing content authored by others." *Ethics and Information Technology* (in press 2021), p.1.

[47] Katherine Clayton, Spencer Blair, Jonathan A. Busam, Samuel Forstner, John Glance, Guy Green, Anna Kawata, et al., "Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media," *Political Behavior* 42, no. 4 (2020): 1073-1095; Stephan Lewandowsky, Ullrich KH Ecker, Colleen M Seifert, Norbert Schwarz and John Cook, "Misinformation and its correction: Continued influence and successful debiasing," *Psychological Science in the Public Interest* 13, no. 3 (2012): 106-131; Anne P. Mintz, *Web of deceit: misinformation and manipulation in the age of social media,* (Medford, N.J.: CyberAge Books), 2012.

[48] Vosoughi et al., "The spread of true and false news," 1146-1151, p.1149. See also Prier, "Commanding the trend," 50-85.

[49] Fighting the spread of COVID-19 misinformation | News | Harvard T.H. Chan School of Public Health; see also Daniele Orsoa, Nicola Federicia, Roberto Copettic, Luigi Vetrugnoa, and Tiziana Bovea, "Infodemic and the spread of fake news in the COVID-19-era," *European Journal of Emergency Medicine* 27, no. 5 (2020): 327-328.

[50] How to fight an infodemic - The Lancet.

Facebook began flagging 'disputed content'[51] (see Appendix 3) and from March, 2019 has required ID verification from those seeking to post or otherwise promote political content.[52]

To help combat the proliferation of fake news it is important to understand that sharing can be a kind of pointing: "Look at that!". It acts as a conductor of one's friends' attention flow: "Have you seen this? And what about this?".[53] Indeed, Pennycook and Rand found that people are more likely to share (and so spread) fake news because they are distracted and lazy (their term), rather than biased towards the viewpoint being shared,[54] and Blaine and Boyer report that rumours perceived to be threatening are more likely to attract attention and be shared.[55] But even if sharing an eye-catching news item, irrespective of its truth-value, is intended as the online equivalent of pointing something out to friends, there are epistemic risks involved. Pennycook et al., for example, found that reading a fake news headline only once is sufficient to increase its perceived accuracy,[56] while repeating it increases the likelihood that it will be believed.[57] Bastick likewise reports that exposure to fake news for less than five minutes is sufficient to unconsciously modify behaviour in a manner that aligns with the fake news story (e.g., exposure to the false story that intelligence is linked to speed produced faster finger tapping in participants in the absence of a reported awareness (on their part) of a change in their behaviour).[58] Feezell (2018) found that users who were exposed to political information on Facebook exhibited increased levels of issue salience, particularly among those with low political interest,[59] whereas Effron and Raj found that people who repeatedly encountered the same misinformation online were not only less ethically concerned about sharing it, irrespective of whether they believed the content to be true, but more likely to express approval of it online.[60] Effron and Raj speculate that repeating blatant misinformation is likely to lead to reduced moral condemnation – both for the act of sharing and the content – because the repeated exposure makes it feel intuitively true. Moreover, De Keersmaecker and Roets identified a *perseverance effect*[61] in relation to fake news whereby the initial influence of incorrect information was not simply undone by pointing out that it was in fact false; rather, belief in the content persevered beyond its revealed falsehood, particularly in those with lower cognitive abilities.

In contrast, *drawing attention* to the potential dubious nature of the content – or what Pennycook et al. refer to as an *accuracy nudge*[62] – is purportedly an effective way to combat mis- or disinformation,[63] as is increasing one's

---

[51] https://www.usatoday.com/story/tech/news/2017/03/06/facebook-begins-flagging-disputed-fake-news/98804948/; see also Petros Iosifidis, and Nicholas Nicoli, "The battle to end fake news: A qualitative content analysis of Facebook announcements on how it combats disinformation," *The International Communication Gazette* 82, no. 1 (2020): 60-81.

[52] Pardis Pourghomi, Milan Dordevic, and Fadi Safieddine, "Facebook Fake Profile Identification: Technical and Ethical Considerations," *International Journal of Pervasive Computing and Communications* 16, no. 1 (2020): 1-9.

[53] Lavinia Marin, "Sharing (mis)information on social networking sites. An exploration of the norms for distributing content authored by others," *Ethics and Information Technology* (in press, 2021).

[54] Gordon Pennycook, and David G. Rand, "Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning," *Cognition* 188 (2019): 39-50.

[55] Timothy Blaine, and Pascal Boyer, "Origins of sinister rumors: A preference for threat-related material in the supply and demand of information," *Evolution and Human Behavior* 39, no. 1 (2018): 6-75.

[56] Gordon Pennycook, Tyrone D. Cannon, and David G. Rand, "Prior exposure increases perceived accuracy of fake news," *Journal of Experimental Psychology: General* 147, no. 12 (2018): 1865-1880.

[57] Lisa K Fazio, Nadia M. Brashier, B. Keith Payne, and Elizabeth J. Marsh, "Knowledge does not protect against illusory truth," *Journal of Experimental Psychology: General* 144, no. 5 (2015): 993-1002.

[58] Zach Bastick, "Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation," *Computers in Human Behavior*, 116 (in press 2021): 1-12.

[59] Jessica T. Feezell, "Agenda setting through social media: The importance of incidental news exposure and social filtering in the digital era," *Political Research Quarterly* 71, no. 2 (2018): 482-494.

[60] Daniel A Effron, and Medha Raj, "Misinformation and morality: Encountering fake-news headlines makes them seem less unethical to publish and share," *Psychological Science* 31, no. 1 (2020): 75-87.

[61] Jonas De Keersmaecker, and Arne Roets, "'Fake news': Incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions," *Intelligence* 65 (2017): 107-110. See also Lee Ross, Mark R Lepper, and Michael Hubbard, "Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm," *Journal of Personality and Social Psychology* 32, no. 5 (1975): 880-892.

[62] Gordon Pennycook, Ziv Epstein, Mohsen Mosleh, Antonio A. Arechar, Dean Eckles, and David G. Rand, "Shifting attention to accuracy can reduce misinformation online," *Nature* (2021, in press): 1-6.

[63] Gordon Pennycook, Jonathon McPhetres, Yunhao Zhang, Jacjson G. Lu, and David G. Rand, "Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy nudge intervention," *Psychological Science* (2021, in press). Moreover, according to Tucker, Facebook's attempts at combating fake news have been effective at reducing the amount of fake or otherwise dubious content shared on Facebook by as much as 75%; see https://mitsloan.mit.edu/ideas-made-to-matter/mit-sloan-research-about-social-media-misinformation-and-elections.

information literacy[64] and being a more analytic thinker.[65] It is worth noting, however, that *if* sharing is intended simply as a way of connecting with one's friends, the truth-value of the shared post need not be its most important asset; nor, from Facebook's perspective, is it a requirement for *prosumerism*.[66]

The Facebook filter bubble[67] means that news editors are not the gatekeepers and agenda setters of prominent news stories that appear on Facebook. Instead, these decisions are more and more a consequence of user engagement (e.g., liking and sharing)[68] and the algorithms that analyse this engagement, neither of which may be focused on, nor concerned with, the direct consumption of news. Filtering algorithms often exert their power (their governance) by either prioritising or even over-emphasising certain content (e.g., news), or censoring it, so that what is fed to the user better aligns with their interests (a kind of *'if you liked that, you are likely to enjoy this'* principle).[69] For this reason, the content is more likely to compete effectively in the *market of attention*.[70]

Thorson et al. argue that an individual's preferences and choices cannot be separated from the algorithms that respond to them.[71] It is not *simply* that the Facebook algorithms shape personal interests; rather, these interests already exist and dictate how the algorithms respond. In other words, initial interest affects algorithm selection which affects content exposure, which likely refines or subtly alters (even if only to reinforce and make more entrenched) these interests. For this reason, Geschke et al. conceive of a *triple-filter bubble* operating in Facebook.[72] The triple-filter comprises technology (e.g., algorithms), social influences (e.g., societal engagement on the platform) *and* individual biases (e.g., confirmatory bias). As such, even in the absence of technology and with limited social engagement, Geschke et al. conjecture, a filter may still operate at an individual level to shape the information sought and retained (such as a bias in one's reasoning, whereby one seeks only information that confirms one's beliefs and disregards that which challenges them).

Facebook's algorithms make editorial decisions about when and where content is made available to the user. The criteria of news editors have therefore been supplanted on social media platforms by the criteria of those creating Facebook's algorithms. Consequently, even though editorial decisions are automated, and not therefore directly implemented by humans, as the criteria *is by human-design in the first place*, there is a risk (a likelihood, in fact) that these criteria will be subject to the designers' biases. In the case of a company like Facebook, it is hardly contentious to suggest that what is driving the design of the algorithms, and the functions they perform (i.e., what content they prioritise, and when and how it is presented), is Facebook's commercial interests.[73] This may, of course, include *social responsibility,* which is also in the commercial interests of Facebook to embrace, or at least appear to be embracing.

---

[64] Mo Jones-Jang, Tara Mortensen, and Jing-Jing Lui, "Does Media Literacy Help Identification of Fake News? Information Literacy Helps, but Other Literacies Don't," *American Behavioral Scientist* 65, no. 2 (2019): 1-18.

[65] Gordon Pennycook, and David G. Rand, "Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning," *Cognition* 188 (2019): 39-50.

[66] While it may not be a *requirement*, one could nevertheless conjecture that it is unlikely that Facebook could maintain its success if it were populated largely by those spreading mis- or disinformation.

[67] The term filter bubble refers to a situation in which an Internet user encounters only information and opinions that conform to and reinforce their own beliefs, caused by algorithms that personalise an individual's online experience. filter bubble definition - Bing.

[68] Michael A. DeVito, "From Editors to Algorithms," *Digital Journalism* 5, no. 6 (2017): 753-773; Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petronic, Antonio Scala, Guido Caldarelli et al., "The spreading of misinformation online," *PNAS* 113, no. 3 (2016): 554-559.

[69] Michael J. Mazarr, Ryan Bauer, Abigail Casey, Sarah Heintz, and Luke J. Matthews, "The Emerging Risk of Virtual Societal Warfare: Social Manipulation in a Changing Information Environment," Santa Monica, CA: RAND Corporation, 2019. https://www.rand.org/pubs/research_reports/RR2714.html.

[70] Nicholas Diakopoulos, "Algorithmic accountability reporting: on the investigation of black boxes," *Columbia Journalism School. Tow Center for Digital Journalism*, 2014: 1-37 Algorithmic Accountability Reporting: On the Investigation of Black Boxes | Tow Center (columbia.edu).

[71] Kjerstin Thorson, Kelley Cotter, Mel Medeiros, and Chankyung Pak, "Algorithmic inference, political interest, and exposure to news and politics on Facebook, Information," *Information Communication & Society* 24, no. 2 (2021): 183-200. DOI: 10.1080/1369118X.2019.1642934.

[72] Daniel Geschke, Jan Lorenz, and Peter Holtz, "The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers," *British Journal of Social Psychology* 58, no. 1 (2018): 129-149. See also Christophe Marret, "The relative impact of the Facebook's algorithm on the creation of social bubbles," (2020) (PDF) The relative impact of the Facebook's algorithm on the creation of social bubbles (researchgate.net).

[73] Presuel, Rodrigo, and Sierra, "Algorithms and the News," 261-285.

The fact that Facebook algorithms, to all intents and purposes, make editorial decisions is particularly striking given DeMarzo et al. and Enke and Zimmermann's findings that social media audiences may have difficulty distinguishing between genuine information and fabricated stories, especially if the misinformation is received from multiple sources.[74] Horwitz and Seetharaman note that a user who likes, shares or comments on 1,500 pieces of content has more influence on the platform and its algorithms than one who interacts with just 15 posts.[75] This bias in Facebook's engagement-based metrics allows "super-sharers" or "influencers"[76] to "drown out" less-active users. They also report that, by Facebook's own reckoning, accounts with hyperactive engagement were far more partisan on average than normal Facebook users. They were also more likely to behave suspiciously (i.e., sometimes active on the platform for as much as 20 hours a day, which suggests either people working in shifts or using bots). Given this, not only is it possible to elevate minor stories to 'headline' status but also to create news stories through targeted mis- and disinformation campaigns by manipulating Facebook's system of news selection, whereby stories, if not curbed, can and do spread exponentially and lead to polarisation.[77] It is for this reason that Helberger describes social media platforms as active political actors capable of wielding considerable political clout[78] (or at least entities that exhibit this *potential*).

To help counter the effect of mis- and disinformation, Giachanou et al. are developing a means of analysing personality traits and linguistic patterns to differentiate between those who are likely to post refuting evidence to combat a fake news story (known as Checkers) and those more likely to help disseminate it (Spreaders).[79] It is also of note that natural language algorithms are being developed to distinguish between fake and genuine news

---

[74] DeMarzo, Peter M., Dimitri Vayanos, and Jeffrey Zwiebel, "Persuasion bias, social influence, and unidimensional opinions," *The Quarterly Journal of Economics* 118, no. 3 (2003): 909-968; Enke, Benjamin, and Florian Zimmermann, (2019). Correlation neglect in belief formation. *The Review of Economic Studies*, 86, no. 1 (2019): 313-332. Each cited in Nicolas M Anspach, and Taylor N. Carlson, "What to believe? Social media commentary and belief in misinformation," *Political Behavior* 42, no. 3 (2020): 697-718.

[75] Horwitz and Seetharaman. "Facebook Executives Shut Down Efforts."

[76] "Influencers" are individuals or small groups that exercise a disproportionate amount of influence over the news flow; see Michael J Mazarr, Ryan Bauer, Abigail Casey, Sarah Heintz, and Luke J. Matthews, "The Emerging Risk of Virtual Societal Warfare: Social Manipulation in a Changing Information Environment. Santa Monica, CA: RAND Corporation, 2019. https://www.rand.org/pubs/research_reports/RR2714.html; also Richard Fletcher, Alessio Cornia, Lucas Graves, and Rasmus Kleis Neilsen, *Measuring the Reach of "Fake News" and Online Disinformation in Europe* (Oxford: University of Oxford), 2018.

[77] Uthsav Chitra, and Christopher Musco, "Analyzing the Impact of Filter Bubbles on Social Network Polarization," WSDM '20 Proceedings of the 13th International Conference on Web Search and Data Mining. (January 2020): 115-123. https://doi.org/10.1145/3336191.3371825.

[78] Natali Helberger, "The Political Power of Platforms: How Current Attempts to Regulate Misinformation Amplify Opinion Power," *Digital Journalism* 8, no. 6 (2020): 842-854.

[79] Anastasia Giachanou, Esteban Ríssola, Bilal Ghanem, Fabio Crestani, and Paolo Rosso, "The Role of Personality and Linguistic Patterns in Discriminating Between Fake News Spreaders and Fact Checkers. International Conference on Applications of Natural Language Systems. NLDB 2020: Natural Language Processing and Information Systems (2020): 181-192; Jennifer Golbeck, Matthew Mauriello, Brooke Auxier, Keval H. Bhanushali, Christopher Bonk, Mohamed Amine Bouzaghrane, et al., "Fake News vs Satire: A Dataset and Analysis." WebSci'18, Amsterdam, Netherlands: Proceedings of the 10th ACM Conference on Web Science (27-30 May 2018): 17-21; Benjamin D. Horne, Sibel Adah, "This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News," The Workshops of the Eleventh International AAAI Conference on Web and Social Media AAAI Technical Report, News and Public Opinion (2017): 1-9. Sehgal et al. also describe how hyperlinks shared by certain Twitter users are being used, in what appears to be a coordinated effort, to spread and amplify misinformation (see Vibhor Sehgal, Ankit Peshin, Sadia Afroz, and Hany Farid, "Mutual hyperlinking among misinformation peddlers." arXiv:2104.11694v1 [cs.SI] 20 Apr 2021).

content,[80] as well as fake news and satire.[81] In addition, eye tracking technology is being developed that is capable of measuring differences in how we track genuine and fake news content.[82]

News consumption, therefore, becomes characterised by selective exposure, sometimes within sets of highly polarised community structures that act as segregated echo chambers.[83] As a consequence:

> The credibility of a news source is an extension of the perceived credibility of the information it provides, which in turn, is driven by how that information matches up with the consumer's partisan predispositions. As media choice siphons casual news viewers away from the regular news audience, partisan news providers become more permanent fixtures in a for-profit news market.[84]

In keeping with Kelly's view, research has found that individuals are more likely to accept information that is consistent with their existing beliefs than information that contradicts them, even when the information is factually inaccurate or otherwise misleading.[85] They are also more likely to circulate false information than to correct misinformation.[86]

Facebook's filter bubble is purportedly intended to help users access content, including news, that matches their interests.[87] These interests are, in turn, influenced by the users' social connections which Facebook 'helps' cultivate. A *charitable* interpretation of the filter bubble is that it *relegates* rather than restricts access to alternative viewpoints and any exposure to fake news is an *unintended* consequence of this process. Whether one agrees with this interpretation or not, Anspach and Carlson (2020) have identified a further potential source of mis- or disinformation that may act independently of the news item's content.[88]

---

[80] Michał Choraś, Konstantinos Demestichas, Agata Giełczyk, Álvaro Herrero, Paweł Ksieniewicz, Konstantina Remoundou, Daniel Urda, Michał Woźniak, "Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study," Applied Soft Computing Journal, 101 (in press 2021); Georgios Gravanis, Athena Vakali , Konstantinos Diamantaras, and Panagiotis Karadais, "Behind the cues: A benchmarking study for fake news detection," Expert Systems and Applications 128 (2019): 201-213; Ray Oshikawa, Jing Qian, William Yang Wang, "A Survey on Natural Language Processing for Fake News Detection," Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020), Marseille, 11–16 May 2020: 6086–6093; David Shariatmadar, "Could language be the key to detecting fake news?" The Guardian, 2 September 2019. Retrieved 19 March 2021 from: Could language be the key to detecting fake news? | Language | The Guardian; Fatemeh Torabi Asr, "The language gives it away: How an algorithm can help us detect fake news," The conversation, 5 August 2019. The language gives it away: How an algorithm can help us detect fake news (theconversation.com) 19 March 2021; Svitlana Volkova, Kyle Shaffer, Jin Yea Jang, and Nathan Hodas, "Separating Facts from Fiction: Linguistic Models to Classify Suspicious and Trusted News Posts on Twitter," Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Short Papers), 647–653 Vancouver, Canada, (30 July – 4 August 2017). 647-653.

[81] Or Levi, Pedram Hosseini, Mona Diab, and David Broniatowski, "Identifying Nuances in Fake News vs. Satire: Using Semantic and Linguistic Cues," Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda, 2019: 31-35. 10.18653/v1/D19-5004. See also Xichen Zhang, and Ali A. Ghorbani, "An overview of online fake news: Characterization, detection, and discussion," Information Processing and Management 57, no. 2 (2020): 1-26 for a detailed fake news typology and detection overview.

[82] Omer Sümer, Efe Bozkir, Thomas Kübler, Sven Grüner, Sonja Utz, Enkelejda Kasneci, "Fake News Perception: An Eye Movement Dataset on the Perceived Believability of News Stories," *Data in Brief* (2021). https://doi.org/10.1016/j.dib.2021.106909.

[83] An environment in which a person encounters only beliefs or opinions that coincide with their own, so that their existing views are reinforced, and alternative ideas are not considered. echo chamber definition - Bing. See also Matteo Cinelli, Mauro Conti, Livio Finos, Francesco Grisolia, Petra K Novak, Antonio Peruzzi, Maurizio Tesconi, Fabiana Zollo and Walter Quattrociocchi, "(Mis)Information Operations," *Journal of Information Warfare* 18, no. 3 (2019): 83-98.

[84] Dimitri Kelly, "Evaluating the news: (Mis)perceptions of objectivity and credibility," *Political Behavior* 41, no. 2 (2019): 445-471, p. 458.

[85] Cengiz Erisen, David P. Redlawsk, and Elif Erisen, "Complex thinking as a result of incongruent information exposure," American Politics Research 46, no. 2 (2017): 217-245. https://doi.org/10.1177/1532673X17725864; Kelly R. Garrett, Brain F. Weeks, B.E., and Rachel L. Neo. "Driving a wedge between evidence and beliefs: How online ideological news exposure promotes political misperceptions," *Journal of Computer-Mediated Communication* 21 (2016): 331-348; Lewandowsky, et al., "Misinformation and its correction," 106-131.

[86] Jieun Shin, Lian Jian, Kevin Driscoll, and François Bar, "Political rumoring on Twitter during the 2012 US presidential election: Rumor diffusion and correction," *New Media & Society* 19, no. 8 (2017): 1214-1235.

[87] Nick Heath, "Trapped in a filter bubble: How Google and Facebook only feed you what you want to hear." ZDnet, https://www.zdnet.com/article/trapped-in-a-filter-bubble-how-google-and-facebook-only-feed-you-what-you-want-to-hear/.html. See also https://www.facebook.com/help/964154640320617/control-what-you-see-in-news-feed.

[88] Anspach, and Carlson, "What to believe?" 697-718.

Facebook newsfeeds (like other social media) are often infused with readers' comments. The social commentary may include posts from recognised members of a user's social network (who are viewed as a trusted source, especially if they are an *opinion leader*). Importantly, though, the comments need not accurately reflect the content of the news item. A naïve reader (or *opinion follower*) may therefore be exposed to mis- or disinformation not simply, if at all, because of the newsfeed's content, which may or may not contain accurate information, but because of the accompanying social commentary, especially if it is read with more interest than the news content itself (i.e., beyond the headline). Thus, Facebook (and other social media) "have the potential to facilitate the spread of inaccurate information *even if the linked news sources are accurate.*"[89]

The Facebook "privacy crisis"

In January, 2017 the US Office of the Director of National Intelligence released a report accusing Russia of running a disinformation campaign to influence the US election in favour of Trump. It was later revealed that, in early 2016, engineers at Facebook discovered Russian activity designed to disrupt the presidential election. Later that year the team discovered Facebook accounts linked to Russian hackers messaging journalists with information from stolen emails. Team member Alex Stamos informed the Facebook general counsel but, as Facebook had no policy on disinformation, no action was taken. In December, 2016 Stamos met with Zuckerberg and other Facebook leaders. While some objected that it might make Facebook appear to be siding with Democrats, Zuckerberg allowed Stamos to extend his investigations. During most of 2017, Facebook investigators traced more ad pages and groups to Russia, and on 6 September, 2017 Stamos briefed the board's audit committee. Board member Erskine Bowles reportedly reacted with hostility due to the potential political fallout and a Facebook blog post was released stating only that Russian agents had spent the modest sum of $100,000 on ads. The following day the *New York Times* published an investigation demonstrating that Russian intelligence had created fake accounts to promote stolen Democratic emails. After an infuriated reception from both parties in Congress, Facebook agreed to hand over the Russian posts. After further stalling, Facebook admitted that some 26 million people had seen the Russian posts.[90]

A number of other matters increased privacy concerns. Slurs by President Trump against Muslims, Mexicans and others on Facebook increased pressure for action against him, although the platform's investigators concluded he had not violated its terms of service. The Cambridge Analytica scandal broke in early 2018 and in September Facebook was forced to disclose that hackers had accessed nearly 15 million accounts.[91] After it was revealed that user growth had slowed due to the scandal, Facebook shares plunged 19%.[92] *The New York Times* ran an editorial, "Facebook cannot be trusted", on 17 November, 2018. That month it was revealed that Facebook had violated EU data law by tracking user locations. Similar revelations continued in 2019. The use of Facebook to stream the Christchurch massacre of 15 March, 2019 also attracted criticism.[93]

Zuckerberg testified to US Congress from 10 to 11 April, 2018 acknowledging his company had done too little to prevent Facebook's misuse and stating: "It was my mistake, and I am sorry". Facebook CEO Sheryl Sandberg testified to the Senate Select Committee on Intelligence on 5 September, admitting Facebook had been too slow to respond to the Russian campaign but asserting that Facebook was now "investing heavily in people and technology to keep our communities safe".[94] An attempt by the UK House of Commons Digital, Culture, Media and Sport Committee was less successful in interrogating Zuckerberg, with its report claiming he had demonstrated "contempt" towards UK Parliament by refusing to appear and complaining of Facebook sending witnesses who "could not or chose not to answer many questions".[95]

---

[89] Anspach, and Carlson, "What to believe?" 697-718, p. 698; emphasis added.
[90] Lawrence Trautman, "Governance of the Facebook Privacy Crisis," *Pittsburgh Journal of Technology Law and Policy* 20 (2019-2020): 99-101. https://doi.org/10.5195/tlp.2020.234.
[91] Trautman, Governance of the Facebook Privacy Crisis, 104-105, 110.
[92] Rupert Neate, "Over $119bn wiped off Facebook's market cap after growth shock," *The Guardian*, 27 July 2018: https://www.theguardian.com/technology/2018/jul/26/facebook-market-cap-falls-109bn-dollars-after-growth-shock.
[93] Trautman, Governance of the Facebook Privacy Crisis, 110-111.
[94] Trautman, Governance of the Facebook Privacy Crisis, 96-97.
[95] House of Commons Digital, Culture, Media and Sports Committee, *Disinformation and 'Fake News': Final Report,* 14.

Facebook and reform

Facebook has responded to the privacy crisis by defending its image and protecting the viability of its business model. In his testimony to Congress Zuckerberg pledged a number of reforms, including: restricting third-party access to Facebook user data; discontinuing the company's purchase of data from private data mills; investing in AI detection algorithms; employing thousands of new cyber personnel to prevent the spread of disinformation; requiring developers to get user approval before accessing data; requiring advertisers running political campaigns to confirm their identity and display their ads publicly.[96] Critics argue, however, that Zuckerberg's apology and promise of reform follows a pattern he and his company have engaged in since his university days, when he was reprimanded by the Harvard Administrative Board for non-consensually posting pictures of female Harvard students on his website, *Facemash*. Zuckerberg was forced to apologise but faced no further sanction.[97] Sherman identifies a similar recurring pattern, including in 2011 when FTC charged Facebook with violating user privacy by sharing data with external parties. The company settled and apologised but substantial reform was not undertaken, later opening the door to the misuse of data by Cambridge Analytica. As Sherman points out, the imperative of protecting its model means almost any fine, scandal or negative publicity would be less of a threat to Facebook's profitability then any substantial reform.[98] Critics have been still more scathing regarding Sheryl Sandberg, with a *New York Times* article in November, 2018 contending that she aggressively lobbied against critics, even employing a Republican opposition research firm to attempt to discredit them.[99]

On 6 March, 2019 Zuckerberg posted a blog in which he claimed Facebook was "pivoting to privacy". This included a commitment to end-to-end encryption on apps, to interoperability to allow people to communicate across apps and networks, to end the long-term shortage of data and not to store sensitive data in countries with weak records on privacy.[100] Critics responded by contending that Facebook's history of data misuse did not make such claims credible.[101] Analyst Ben Thompson argued that such a strategy makes perfect sense for Facebook, with Zuckerberg's key points centred around the strategy of owning the "one-to-one private ephemeral space" of personal communication through apps, while providing the public relations benefit of allowing Facebook to spruik its "pivot to privacy". Interoperability, he argued, works in Facebook's favour, allowing Facebook to leverage data from Facebook to provide targeted advertising on privacy focused platforms.[102] Tufekci has also contended that interoperability would serve Facebook's interests in avoiding anti-trust remedies while not impeding the collection of data.[103]

Does Facebook's business model hinder attempts to prevent malign actors manipulating its persuasive techniques and technology?

Facebook is incentivised to remove inauthentic accounts – e.g., opened by fraudulent humans and bots – as a means of maintaining the authenticity, and subsequent marketable value, of its commodities.[104] After all, advertisers are not going to be interested in inaccurate information that leads them to target audiences whose

[96] Len Sherman, Zuckerberg's Broken Promises Show Facebook Is Not Your Friend.
[97] Katherine A. Kaplan, "Facemash Creator Survives Ad Board," *The Harvard Crimson*, 19 November 2003: https://www.thecrimson.com/article/2003/11/19/facemash-creator-survives-ad-board-the/.
[98] Len Sherman, "Zuckerberg's Broken Promises Show Facebook Is Not Your Friend."
[99] Sheera Frenkel et al. "Delay, Deny, Deflect: How Facebook Leaders Leaned Out in Crisis," *New York Times*, 15 November 2018: https://www.nytimes.com/2018/11/14/technology/facebook-data-russia-election-racism.html.
[100] Mark Zuckerberg, "A Privacy Focused Vision for Social Networking," *New York Times*, 6 March 2019: https://www.nytimes.com/2019/03/06/technology/facebook-privacy-blog.html.
[101] David Maher, "Why Zuckerberg's Privacy-Focused Vision for Facebook Is Not Entirely Pro-Privacy," *Fortune*, 7 March 2019: https://fortune.com/2019/03/07/facebook-privacy-zuckerberg-whatsapp/.
[102] Ben Thompson, "Facebook's privacy cake," *Stratechery*, 7 March 2019: https://stratechery.com/2019/facebooks-privacy-cake/.
[103] Zeynep Tufekci, "Zuckerberg's So-Called Shift toward Privacy," *New York Times*, 7 March 2019: https://www.nytimes.com/2019/03/07/opinion/zuckerberg-privacy-facebook.html.
[104] Katharina Krombholz et al., "Fake identities in social media," 175-212.

artificial or otherwise fraudulent members are not going to buy what they are selling.[105] In addition, given that Facebook is in the business of selling information about its users to third-parties, it needs to ensure that only it and no-one else has access to the small data points from which the marketable information is derived. Facebook must therefore protect against infiltration by external agencies for two reasons. First, to be able to assure its users that their personal information is secure, so that they will continue to engage with the platform and provide Facebook with the small data points it requires.[106] Second, to have access to voluminous amounts of user information that would-be competitors do not possess.

Facebook's approach to platform integrity is described by Weedon et al.:

> Facebook is a place for people to communicate and engage authentically, including around political topics. If legitimate voices are being drowned out by fake accounts, authentic conversation becomes very difficult. Facebook's current approach to addressing account integrity focuses on the *authenticity of the accounts* in question and their *behaviors*, not the *content of the material created*.[107]

It is true that it would become harder to have an authentic conversation on a platform increasingly occupied by fake accounts. But if the *content* discussed contains more and more information that is factually incorrect, it would also become increasingly difficult to have a *meaningful* conversation. Therefore, while it is arguably the case that the conversation would be authentic insofar as it occurs between individuals who are genuine platform users, it would nevertheless lack epistemic value because it would fail to *inform* and, instead, only misinform those involved in the discussion.

Facebook's reluctance to target content can be grounded on legitimate concerns about protecting freedom of speech. It is also fair to say, however, given Facebook's business model, that the organisation is not incentivised to weed out mis- or disinformation – quite the reverse, in fact – particularly as such content is shared more frequently than authentic content[108] and, therefore, helps cultivate engagement through 'likes', shares and by posting comments. Facebook knows this, just as it knows about research showing that emotionally salient content is more likely to be viewed and shared than emotionally neutral content; it has, after all, commissioned some of this research itself (recall, for example, the 2012 emotional contagion study)[109] and is happy to publicise these findings. Why? Because it is in Facebook's interest to do so: that is, to inform its customers – advertisers – that it (Facebook) knows how to influence the customer's target audience.[110] This extends to Facebook claiming that, by monitoring posts and photos in real time, it is able to identify when a teenager (for example) feels insecure and worthless – i.e., 'stressed', 'defeated', 'overwhelmed', 'anxious', 'nervous', 'stupid', 'silly', 'useless' and a 'failure' – and in need of a confidence boost, information that would no doubt be of value to advertisers and marketing companies.[111] Natural language algorithms are also being developed that are capable of identifying which group membership is most salient within a person's textual exchange (e.g., mother, conservative, feminist).[112] It is not inconceivable that Facebook has similar capabilities.

---

[105] Pertinent to this point, recent industry feedback regarding advertising on Facebook is that their campaign analytics cannot be relied on. This claim is based on personal correspondence with Dr Violetta Wilk (21 May 2021) who adds: A senior marketing manager reported that they had run FB campaigns. All were extremely well-performing (producing high engagement, great click throughs). On closer inspection of profiles that were engaging with the campaigns, however, it was discovered that many were not authentic, had no or little followers, and were engaging with various FB advertising campaigns regularly. Feedback like this is creating scepticism around advertising on Facebook.

[106] Freeze et al., "Fake claims of fake news," 1-33.

[107] Jen Weedon, William Nuland, and Alex Stamos, "Information operations and Facebook." Facebook, Inc., p.10; emphasis added https://www.mm.dk/wp-content/uploads/2017/05/facebook-and-information-operations-v1.pdf.

[108] Prier, "Commanding the trend," 50-85.

[109] Kramer et al., "Experimental evidence of massive-scale," 8788-8790.

[110] Bernal, "Facebook," 513-530.

[111] Sam Levin, "Facebook told advertisers it can identify teens feeling 'insecure' and worthless," *The Guardian*, May 2 2017. Retrieved 18 March 2021 from: Facebook told advertisers it can identify teens feeling 'insecure' and 'worthless' | Facebook | The Guardian.

[112] Miriam Koschate, Elahe Naserian, Luke Dickens, Avelie Stuart, Alessandra Russo, and Mark Levine. "ASIA: Automated Social Identity Assessment using linguistic style," *Behavior Research Methods* (in press, 2021) https://doi.org/10.3758/s13428-020-01511-3.

Fake news tends to elicit from users precisely the sort of behaviour Facebook wishes them to emulate (i.e., increased 'social' engagement). It is therefore in the interest of Facebook to make such targeted content more readily available to those who seek it out, irrespective of its truth status. But it is also in Facebook's interest, at the very least, to appear to be embracing its social responsibilities, which include not deliberately allowing the dissemination of malign content.

Some final remarks

The Facebook business model is essential to the continued existence of Facebook in its present form. In this respect, Facebook could be regarded as vulnerable. Alternatives models for a media company that allows users to share information on a social platform without the retention of data have been discussed. A subscription-based service has been suggested and a donation-led platform based on the Wikipedia model.[113] The viability of the Facebook model could also be undermined by legislation by national governments prohibiting the harvesting of user data. There is some evidence that alternatives could be attractive to users. A 2015 survey of EU citizens revealed that 63% of citizens did not trust online businesses, that more than half did not like providing personal information in return for free services and 53% did not like the fact Internet companies use their personal information in tailored advertising. A 2018 US survey yielded similar results, as did a 2016 survey in the Asia-Pacific.[114] Facebook in its present form did not exist a relatively short number of years ago and, if its business model is undermined, it is not inconceivable it could cease to exist as the large, profitable, transnational entity it is today.

The retention and use of user data is at the heart of the Facebook business model and essential to the platform's survival. Because of this, there is a fundamental mismatch between how privacy campaigners view privacy issues and how such matters are understood by the company. Privacy campaigners see the retention of data by Facebook as the issue, while Facebook understands privacy to be the use it makes of the data it obtains. Facebook accordingly sees itself as the gatekeeper and arbitrator of privacy, with a right to extract and retain the use of data as it sees fit, as well as determine who and how others make use of it. Given this, it is unsurprising that Facebook is often in non-compliance with laws and agreements. As the UK House of Commons report makes clear, if Facebook had complied with the settlement it reached with FTC in 2011 the Cambridge Analytica scandal would not have been possible.[115] There is no evidence that this pattern by Facebook has essentially changed.[116] It is, however, essential that any response to malign operations on Facebook deal with the platform as it now is. Facebook is unlikely to reform its fundamental business model and efforts to regulate it in ways that would substantially change its model do not appear to be imminent. Any response to potentially malign threats must, accordingly, be based on countering these threats on the platform on which they exist and presently thrive, based on the assumption that Facebook's current business model is likely to continue, along with the societal risks it contains.

Facebook is motivated to remove inauthentic accounts because such accounts have a detrimental effect on its marketable commodity (after all, advertisers are hardly likely to be interested in targeting inauthentic accounts). Where inauthentic accounts are responsible for producing and/or spreading fake news, Facebook can be seen to be reducing mis- and disinformation. But Facebook is now engaged more directly and actively in reducing fake news items on its platform, as we will discuss in Section Four.

---

[113] Christine Fisher, "Wikipedia co-founder wants to give you an alternative to Facebook and Twitter," Engadget, 14 November 2019: https://www.engadget.com/2019-11-14-wtsocial-wikipedia-wales-facebook-alternative.html . Also, in October 2019 Wikipedia co-founder Jimmy Wales launched WT.Social, a social media platform financed by user donations without data collection.
[114] Angel Cuevas et al., Does Facebook Use Sensitive Data for Advertising Purposes?, 62.
[115] House of Commons Digital, Culture, Media and Sports Committee, *Disinformation and 'Fake News': Final Report*, 90.
[116] House of Commons Digital, Culture, Media and Sports Committee, *Disinformation and 'Fake News': Final Report*, 90.

Theme 2: Persuasive Technology

Research Question 2: How does Facebook use technology to persuade its target audience?

To address Research Question 2, it is important to establish first who Facebook considers its target audience to be – i.e., the nature and purpose of this audience for Facebook – and why its members need to be persuaded. One can then examine the audience's relationship to third-party actors and organisations who purchase Facebook's audience-based commodities, and how these entities, with the help of Facebook, influence their target audience.

Who does Facebook consider its target audience to be?

In simple terms, Facebook's audience is the nearly three billion individuals who have a Facebook account and use of its platform. But Facebook's audience should not be thought of merely as *an assembled group of spectators to some public event* (as one Oxford dictionary definition of 'audience' states); nor is it merely *a collection of people who give their attention to something* (as another states). Instead, to be a member of Facebook's audience, a more active form of engagement is required than suggested by either of these definitions. This is because Facebook's audience is treated by the company as a product. Engagement of the 'right' kind is therefore required for users to be considered part of Facebook's audience, for such engagement constitutes a means of Facebook productivity. The interactivity afforded by social media platforms (including Facebook) has enabled the traditionally more passive, pre-social media audience to become an active and engaging community.[117]

The utilisation of users' activities on Facebook has been reconceived by Fumagalli et al. as unpaid *digital labour*.[118] They argue that while much digital labour is performed for fun – as social activity, perhaps borne of a spirit of community – it nevertheless has value, insofar as it has the potential to generate a profit for the host organisation (i.e., Facebook), if managed correctly. The Facebook user, in developing their social network, acts as both the consumer and producer of information. They can thus be thought of as *prosumers*[119] (or, as Bruns refers to them, *produsers*[120]) and the act of 'prosumption' is integral to the creation of Facebook's main product: user data acquired through sustained audience engagement.[121] The latter being a means of continually producing the former.

Facebook offers its service– marketed as a means to facilitate social networking[122] – to its audience of prosumers for free because, in return, it obtains from them a valued commodity: the product of their freely given digital labour, in the form of big data,[123] which is then traded to third-party companies, mostly advertisers.[124] Through continuing acts of prosumption, Facebook is able to obtain and store vast quantities of artificial intelligence about the attitudes and behaviours – e.g., likes/dislikes, social relations – of its billons of users, the use of which goes largely unregulated.[125] As White and Boatwright explain:

---

[117] Rupak Rauniar, Greg Rawski, Jei Yang, and Ben Johnson, "Technology acceptance model (TAM) and social media usage: an empirical study on Facebook," *Journal of Enterprise Information Management* 27, no. 1 (2014): 6-30.

[118] Andrea Fumagalli, Stefano Lucarelli, Elena Musolino, and Giulia Rocchi, "Digital labour in the platform economy: The case of Facebook," *Sustainability* 10, no. 6 (2018): 1757, 1-16.

[119] George Ritzer, and Nathan Jurgenson, "Production, consumption, presumption," *Journal of Consumer Culture* 10, no. 1 (2010): 13-36.

[120] Axel Bruns, "The Future is User-Led: The Path Towards Widespread Produsage." *Proceedings of Perth DAC 2007: The 7th International Digital Arts and Culture Conference,* edited by A. Hutchinson, 68-77. Curtin University of Technology, Western Australia, Perth, 2007.

[121] Fumagalli et al, "Digital labour," 1-16.

[122] As noted by Lilley et al., the Facebook homepage reads: Facebook helps you connect and share with people in your life.

[123] Big data are extremely large data sets – comprising activities, comments (including 'Likes', click streams, events, visits to sites – that are analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions. See Christine L. Borgman, *Big data, little data, no data: Scholarship in the networked world* (Cambridge, Mass: MIT Press), 2015.

[124] Shoshana Zuboff, "Big other: Surveillance capitalism and the prospects of an information civilization," *Journal of Information Technology* 30, no. 1 (2015): 75-89.

[125] Stacy McLachlan, "27 Facebook demographics to inform your strategy in 2021," https://blog.hootsuite.com/facebook-demographics/#:~:text=6.%20Facebook%20reaches%20the%20largest%20number%20of%20users%20aged%2013%2D17andtext=Instagram%20reach%20for%20the%2013,teens%20compared%20to%20last%20year.

Facebook does not sell the small data points, which [it] refers to as 'personal information,' since they are of no value. However, small data (personal information), when aggregated, produce artificial intelligence worth billions of dollars. Aggregated small data become big data, which can be further analyzed to produce algorithms that reveal predictive patterns of behaviors, values, and attitudes.[126]

To obtain these small data points (personal information) and convert them into (continually updated) user profiles of interest to advertisers in the form of big data and predictive algorithms,[127] Facebook must keep its audience engaged *as* prosumers. Thus, while Facebook is more than one thing to its audience,[128] anyone who uses Facebook can safely assume that, *to the company*, all members of the Facebook 'audience' are nominally categorised as the same thing: a source of valuable data.[129]
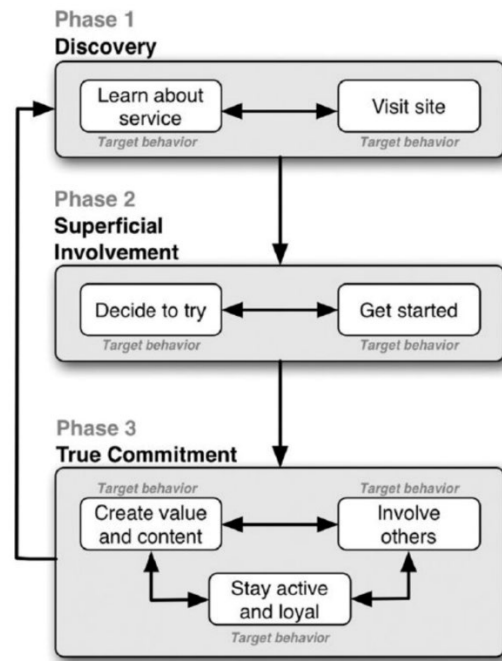

Figure 7: Fogg and Eckles' behavioural chain of committed and therefore valued Facebook use
The loop (phase 3 back to phase 1) indicates the need to keep learning about additional
Facebook services and introduce others to these or Facebook generally.

To maximise the production of data, Facebook users are encouraged to increase their commitment to the platform through a process Fogg and Eckles refer to as a *behavioural chain* (see Figure 7),[130] according to which users proceed through the initial stages of discovery and superficial involvement before truly committing to their role as prosumers (the end-state intended by the Facebook design). As evidence that they are truly committed, Facebook users must create value (through engagement), involve others, and stay active and loyal. To help achieve this, they

---

[126] Candace L. White, and Brandon Boatwright, "Social media ethics in the data economy: Issues of social responsibility for using Facebook for public relations," Public Relations Review 46, no. 5 (2020): 1-7., p.2.
[127] Elinor Carmi, "Rhythmedia: A study of Facebook immune system," *Theory, Culture and Society* 37, no. 5 (2020): 119-138.
[128] Lila MacLellan, "There are only four types of Facebook users, researchers have found," *Quartz*, 13 July 2017 https://qz.com/1026914/the-four-types-of-facebook-users-relationship-builders-window-shoppers-town-criers-and-selfies/.
[129] In the case of Facebook, this is typically language data in the form of posts, responses, messages; multimodal data in the form of photos and videos; and behavioural data in the form of activities and interactions (e.g., A is now friends with B, A liked B's post, A checked in at Sydney Airport). Moreover, whether Facebook sells data directly to advertisers, rather than access to its platform's users, is a contentious issue; see, for example, https://www.bbc.com/news/technology-46618582.
[130] Brian J. Fogg, and Dean Eckles. "The behavior chain for online participation: How successful Web services structure persuasion." Paper presented at the International Conference on Persuasive Technology, Berlin, Heidelberg, 2007. See also Soroush Vosoughi, Deb Roy, and Sinan Aral, "The spread of true and false news online," *Science* 359, no. 6380 (2018): 1146-1151.

must (i.e., are encouraged to) keep their personal profiles up-to-date,[131] invite friends, respond to others' contributions and return to the platform often.[132] Facebook facilitates these steps by making engagement quick and simple (e.g., often one-click requirements) and 'keeps track' of user actions, such as signalling preferences – say, in the form of 'likes' and 'dislikes' – or adding connections (i.e., allowing access to one's password protected email list), friending/unfriending and so on.[133] But as van Dijck notes:

> "Keeping track" is a euphemism for data mining. In fact, [Facebook] heavily invests in tracking, interpreting, repurposing, and selling information generated by its users. What are patterns of likeability? Which people who like X also like Y? Which circles of friends are sensitive to which preferences or joint tastes? Whereas "making connections" may be the main goal for users to share information, for [Facebook,] the main goal is to extract valuable collective and personalised information out of a data flow of extracted "likes," "pokes," "dislikes" and so on. Users often leave behind a trail of inconspicuous information that becomes a prime asset for application programmers.[134]

Facebook also 'keeps track' of users *and non-users* by creating *shadow profiles*. Shadow profiles contain information about an individual that Facebook has acquired indirectly through accessing another Facebook user's account. To illustrate, suppose user A allows Facebook to access her personal email list. On this list is an email address of user B that is different to the email user B has given to Facebook (say, when registering). Facebook will store user B's second email address in a shadow profile. This same principle applies even in the case information about individuals who are not registered with Facebook. Information about individual C (a non-Facebook user) gleaned by accessing other users' accounts is stored in a shadow profile for potential future reference, should individual C ever join Facebook.[135]

The long-term viability of Facebook therefore depends on maintaining a delicate balance between *attracting* an audience and *exploiting* its natural resources (i.e., individual member's attitudes and behaviour) while maintaining credibility.[136] To do this, Facebook acts as an audience engagement tool (see Figure 8),[137] accommodating and entertaining users through the social connections they develop via its platform. As an audience engagement tool, Facebook seeks to influence individuals to participate further and more often, and in the *right way*, on its online platform and, in so doing, provide more of the source material (the small data points) on which its most prized commodities – big data and predictive algorithms – depend.

---

[131] Katharina Krombholz, Dieter Merkl, and Edgar Weippl, "Fake identities in social media: A case study on the sustainability of the Facebook business model," *Journal of Service Science Research* 4, no. 2 (2012): 175-212.

[132] Brian J. Fogg, and Daisuke Iizawa, "Online persuasion in Facebook and Mixi: A cross-cultural comparison," Paper presented at the International Conference on Persuasive Technology, Berlin, Heidelberg, 2008.

[133] José van Dijck, "Facebook as a tool for producing sociality and connectivity." *Television & New Media* 13, no. 2 (2012): 160-176.

[134] José van Dijck, "Facebook as a tool for producing sociality and connectivity," 160-176, p.168.

[135] Shadow profiles - Facebook knows about you, even if you're not on Facebook (theconversation.com).

[136] Brian Cugelman, Mike Thelwall, and Phil Dawes, "Website credibility, active trust and behavioural intent," Paper presented at the International Conference on Persuasive Technology, Berlin, Heidelberg, 2008. Cugelman distinguishes between presumed credibility (a pre-use view), reputed credibility (which is shaped by third-parties), surface credibility (which is acquired after superficial engagement), and experienced credibility (something that can only be acquired after repeated use).

[137] Virginia H. Balfour, "Likes, comments, action! An examination of the Facebook audience engagement strategies used by strategic impact documentary," *Media International Australia* 176, no. 1 (2020): 34-51.
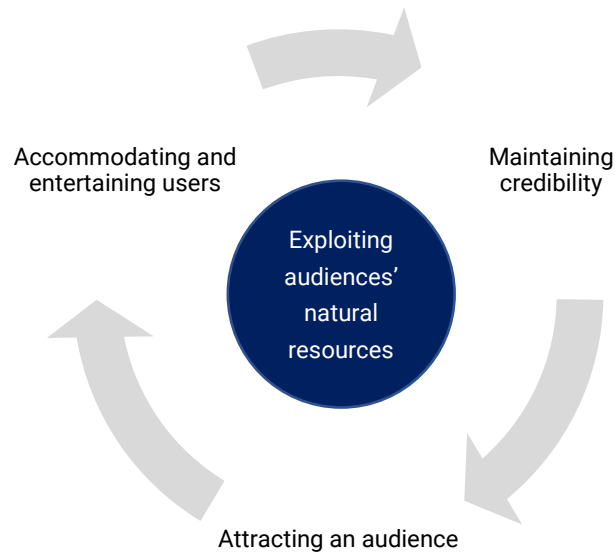
Figure 8: Facebook as an audience engagement tool
Attracting an audience by accommodating and entertaining them while maintaining credibility, which
further attracts an audience. This is balanced against exploiting the audience's natural resources.

2.2 How does Facebook influence its audience to participate in the 'right' way on its platform?

Lilley at al. correctly note that Facebook users are not *forced* to engage with the social networking site, nor participate as prosumers on a platform that intentionally exposes them and their social networks to commercialisation.[138] Instead, they do so willingly – but also, at times, unwittingly[139] – and continue to do so because their behaviour and, ultimately, their willingness to engage with the platform is, to some degree, engineered by Facebook.

---

[138] Stephen Lilley, Frances Grodzinsky, and Andra Gumbus, "Revealing the commercialized and compliant Facebook user," *Journal of Information, Communication and Ethics in Society* 10, no. 2 (2012): 82-92.
[139] Once a user is logged in, many of their Internet interactions are accessible by Facebook. Consent is almost certainly not given in these circumstances. See Adi Robertson, "Sites Could Be Liable for Helping Facebook Secretly Track Your Web Browsing, Says EU Court." 29 July 2019. https://www.theverge.com/2019/7/29/8934924/facebook-like-buttons-court-justice-eu-ruling-privacy-data-protection-tracking.

Figure 9: Cultivating a trusted and social Facebook environment

The most commonly reported usage of Facebook is to stay in touch with friends[140] or, more formally, to engage in *interpersonal interconnectivi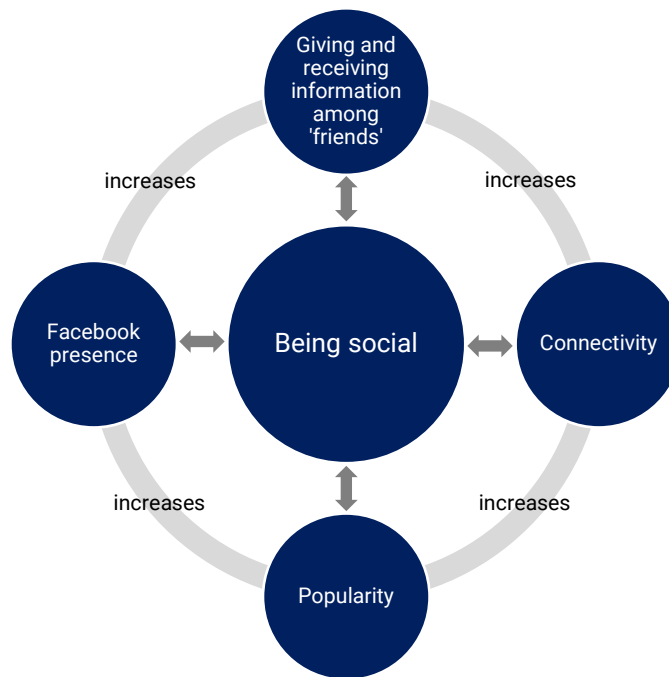ty*[141] or *relation maintenance*.[142] In support of this claim, Wilson et al. found a strong offline-to-online trend, whereby interpersonal relationships established offline were extended to include the online arena.[143] Interpersonal interconnectivity aligns with one of the two motivations for using Facebook reported by Nadkarni and Hofmann, the other being the need for self-presentation: both of which speak to the sociality of Facebook.[144] Similarly, You et al. claim that Facebook (and other social media) facilitate *culture sharing*, which they define as "the exchanges or mutual exposures of preferred lifestyles via social ties between users from different cultural backgrounds."[145] Carmi, however, offers a word of caution when considering the *social* aspect of social media sites like Facebook, by stating that the way we come to understand what it is to be 'social' on these platforms "is influenced by their own definitions of what it means to be social".[146]

A key aspect of being social on Facebook is sharing and building one's social network of 'friends', the former being a way to achieve the latter. This network evolves into a trusted source in terms of the giving and receiving of information.[147] Thus, users tend to trust their personal information to Facebook because they are often invited to

---

[140] Khe Foon Hew, "Students' and teachers' use of Facebook," *Computers in Human Behavior* 27, no. 2 (2011): 662-76; Ohbyung Kwon, and Yixing Wen, "An empirical study of the factors affecting social network service use," *Computers in Human Behavior* 26, no. 2 (2010): 254-63.

[141] Christy Cheung, Pui-Yee Chiu, and Matthew K.O. Lee, "Online social networks: Why do students use Facebook?" *Computers in Human Behavior* 27 (2011): 1337-1343.

[142] Nazan Dogruer, Ipek Menevis, and Ramadan Eyyam, "What is the motivation for using Facebook?" *Procedia Social and Behavioral Sciences* 15 (2011): 2642-2646.

[143] Robert E. Wilson, Samuel D Gosling, and Lindsay T Graham, "A review of Facebook research in the social sciences," *Perspectives on Psychological Science* 7, no. 3 (2012): 203-20.

[144] Ashwini Nadkarni, and Stefan G Hofmann, "Why do people use Facebook?" *Personality and Individual Differences* 52, no. 3 (2012): 243-249.

[145] Quenzeng You, Darío García- García, Manohar Paluri, Jiebo Luo, and Jungseock Joo, "Cultural Diffusion and Trends in Facebook Photographs," Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017), page 347 Cultural Diffusion and Trends in Facebook Photographs - Facebook Research (fb.com).

[146] Carmi, "Rhythmedia," 119-138. p.121.

[147] Samantha Henderson, and Michael Gilding "'I've never clicked this much with anyone in my life': Trust and hyperpersonal communication in online friendships," *New Media & Society* 6, no. 4 (2004): 487-506.

join by a 'friend' who they know as part of their offline circle of friends and acquaintances.[148] Moreover, an individual's 'popularity' on the platform is rooted in the connections they establish as part of their social network and, so, connectivity built on trust translates into a quantifiable commodity.[149] Users can enhance the *saliency* of their 'popularity' through increased social presence which one could calculate in terms of the number of posts, 'likes', images and 'friends' they have[150] (see Figure 9). But as van Dijck notes, popularity "is created through engineered algorithms that prompt users to rank things, ideas, or people in relation to other things, ideas, or people. [And while] popularity has no relation to values like truth, trust, objectivity, or quality, it is often equated to these values."[151]. Because of this, for some (perhaps many), the risk of limiting access to personal information and losing one's salient social presence (and subsequent popularity) is perceived as greater than the risk brought about from disclosure.[152] That said, while the motivation to maintain social connections online is still the most popular reason for using Facebook,[153] more recently, Hong and Oh reported that Facebook's popularity among the young is beginning to wane; to be replaced by other social media sites or, in some cases, as a reaction against social media in general.[154]

Facebook's algorithmic-based assessment of individual preferences is used to determine the ease with which a user can access content on their Facebook page (e.g., certain newsfeeds and advertisements rather than others) by selecting – prioritising – what is made more readily available to view (see Appendix 4 for a timeline of algorithm development).[155] Waldman reports that Facebook's News Feed algorithm was tweaked during system improvements between 2015-2016 to assign greater weight (i.e., preference) to posts and interactions from friends and family.[156] This prioritised material is often shared by the user (a trusted source) with others in their social network,[157] and likely gains traction among the group of 'friends' as a news story or item of interest because of this.[158] Such cultivated and trusted sharing can, however, make users more susceptible to mis- or disinformation, more commonly known as fake news.[159] To illustrate: The findings of Facebook's (infamous) 2012 *emotional*

---

[148] van Dijck, "Facebook as a tool," 160-176.

[149] van Dijck, "Facebook as a tool," 160-176; Benjamin Grosser, "What do metrics want? How quantification prescribes social interaction on Facebook," *Computational Culture*, no. 4 (2014): 1-41.

[150] Emily Christofides, Amy Muise, and Serge Desmarais, "Information disclosure and control on Facebook: Are they two sides of the same coin or two different processes?" *CyberPsychology and Behavior* 12, no. 3 (2009): 341-345. Tom Tong et al. found, however, that too many 'friends' has a negative impact on trust others have in the authenticity of the profile.

[151] van Dijck, "Facebook as a tool," 160-176. p. 170.

[152] Christofides et al. "Information disclosure and control," 341-345.

[153] (GlobalWebIndex 2017) The 10 Top Reasons Why We Use Social Networks [Updated] (wersm.com) (Accessed 21 March 2021).

[154] Seoyeon Hong, and Sookwang Klive Oh, "Why People Don't Use Facebook Anymore? An Investigation into the Relationship Between the Big Five Personality Traits and the Motivation to Leave Facebook." *Frontiers in Psychology* 11, article 1487 (2020): 1-9. See also Growing social media backlash among young people, survey shows | Social media | The Guardian and Why we millennials are happy to be free of social media tyranny | Social media | The Guardian. For other reasons why people do not use Facebook, see Anas Aloudat, Ons Al-Shamaileh, and Katina Michael, "Why some people do not use Facebook?: *Social Network Analysis and Mining* 9, no. 1 (article 19) (2019):1-14.

[155] See also Balancing Content from Friends and Pages - About Facebook (fb.com) and https://about.fb.com/news/2016/06/news-feed-fyi-helping-make-sure-you-dont-miss-stories-from-friends/.

[156] Ari E. Waldman, "Privacy, sharing, and trust: The Facebook study," *Case Western Reserve Law Review* 67, no. 1 (2016b): 193-234. See also Callum Borchers, "Facebook invites you to live in a bubble where you are always right," *The Washington Post*. https://www.washingtonpost.com/news/the-fix/wp/2018/01/14/facebook-invites-you-to-live-in-a-bubble-where-you-are-always-right/.html.

[157] Paul Bernal, "Facebook: Why Facebook makes the fake news problem inevitable," *Northern Ireland Legal Quarterly* 69 (2018): 513-530. See also Aris Anagnostopoulos, Ravi Kumar, and Mohammad Mahdian, "Influence and Correlation in Social Networks," KDD '08: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, 24–27 August 2008, Las Vegas, Nevada, USA, 7-15 https://doi.org/10.1145/1401890.1401897 for discussion on social influence.

[158] Nicolas M. Anspach, "The new personal influence: How our Facebook friends influence the news we read," *Political Communication* 34, no. 4 (2017): 590-606.

[159] David M.J. Lazer, Matthew A Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, *et al.*, "The science of fake news," *Science* 359, no. 6380 (2018): 1094-1096. See also Vosoughi, et al., "The spread of true and false news," 1146-1151.

contagion study[160] reveal that emotional content yields higher levels of engagement in terms of comments and shares compared to emotionally neutral content (e.g., a post about food or interior design).[161] Facebook therefore has an added incentive – in terms of increasing prosumer engagement – to prioritise or privilege news items or other content that will elicit an emotional reaction.[162] Moreover, this incentive remains irrespective of the epistemic value of the content; irrespective, that is, of its truth or falsity.

In 2021, Facebook announced a series of changes to its News Feed *Feed Filter Bar* to facilitate greater user-control over what they see and who can comment on their posts.[163] For example, "Most Recent" allows a user to switch from algorithmically-ranked to chronologically-ranked news items. With "Favourites", a user can select up to 30 friends and/or pages whose content is prioritised or displayed in a separate feed. Other changes include, turning off political ads and temporarily hiding posts from a person, group or page. Users also have greater control over who can comment on their posts. Those permitted to comment can be restricted to only the profiles and pages mentioned in "Who can comment on your post?". This extra control is particularly useful to media companies who are held legally responsible for the content of comments posted on their Facebook page (see Section 4.3 for additional discussion), even when, as used to be the case prior to this change, they had no control over who posted comments on public posts, including removing them.[164]

Does Facebook employ techniques of non-rational persuasion or manipulation?

According to Waldman, the typical Facebook user does not make perfectly rational decisions about sharing information, whether personal or otherwise.[165] Facebook capitalises on this because it is designed to facilitate non-rational sharing (see Appendix 5 for further details on the continuum of influence of which non-rational persuasion is a part).[166] It nudges us to share, Waldman claims, by *scratching its users' social itches*,[167] often through the ease by which we can click 'like' on new content and share this with others.[168] For Waldman, Facebook encourages (as in, manipulates) users to share personal information by designing its platform to cultivate and promote trust as a resource among its members, insofar as it is a means for them to acquire social capital. This could be said of any social situation where one seeks to increase one's social capital, of course; so, certainly, what is described is not unique to Facebook. Nevertheless, Waldman is critical of Facebook's design tactics which he claims:

> Leverage the trust we have in our friends to manipulate us into sharing personal information with websites, advertisers, and third party partners we've never met or heard of. When it does, Facebook crosses the line from… a conduit of social sharing to a manipulative for-profit scheme where users are reduced to the terabytes of data they generate.[169]

---

[160] Adam D.I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *Proceedings of the National Academy of Sciences* 111, no. 24 (2014): 8788-8790. See also Jukka Jouhki, Epp Lauk, Maija Penttinen, Niina Sormanen, and Turo Uskali, "Facebook's Emotional Contagion Experiment as a Challenge to Research Ethics," *Media and Communication* 4, no. 4, 2016: 75-85.

[161] See Karen Nelson-Field et al., "The emotions that drive viral video," *Australasian Marketing Journal* 21, no. 4 (2013): 205-211 The Emotions that Drive Viral Video (sagepub.com) for similar findings, and Melanie Schreiner, et al., "Impact of content characteristics and emotion on behavioral engagement in social media: Literature review and research agenda," *Electronic Commerce Research*. (2019, in press): 1-17. Impact of content characteristics and emotion on behavioral engagement in social media: literature review and research agenda (springer.com) for a detailed overview of recent research.

[162] Carmi, "Rhythmedia," 119-138.

[163] https://about.fb.com/news/2021/03/more-control-and-context-in-news-feed/.

[164] Facebook to restrict comments on public posts (smh.com.au).

[165] Ari E. Waldman, "Manipulating trust on Facebook," *Loyola Consumer Law Review* 29, no. 1 (2016a): 175-198.

[166] See, for example, *dark patterns* https://www.darkpatterns.org/.

[167] A phrase he borrows from James Grimmelmann, "Saving Facebook Iowa law review," *New York Law School*, 94 (2009): 1137-1151.

[168] Lisa Harris, and Charles Dennis, "Engaging customers on Facebook: Challenges for e-retailers," *Journal of Consumer Behaviour* 10, no. 6 (2011): 338-346.

[169] Waldman, "Manipulating trust on Facebook," 175-198. p.177.

Table 1: Cialdini's seven techniques of influence

Technique of influence	People are likely to be persuaded by you if…
Social proof (or consensus)	They believe that lots of other people have or would do the same
Reciprocity	You have helped them previously
Authority	They believe you are an expert or in a position of authority
Liking	They like you, because they perceive you like them, or because you (are perceived to) have things in common.
Commitment and consistency	They believe that what is being asked is consistent with other things they have done or with their values (it demonstrates commitment to these things)
Scarcity	They believe what is being asked will enable them to obtain a unique, rare, or otherwise scarce resource
Unity	They perceive themselves as similar to, or to have something in common with, you

Waldman is clear about the type of influence Facebook exerts. A more charitable interpretation, however, might lead one to conclude that Facebook employs techniques of non-rational persuasion rather than manipulation. To illustrate: In 1984, Cialdini proposed six techniques of influence (although he also uses the term 'persuasion'), adding a seventh more than three decades later [170] (see Table 1). Cialdini's techniques are compatible with the notion of non-rational persuasion and can be used to explain Facebook's success at shaping its users' attitudes and behaviour (see Figure 10), although the line between non-rational persuasion and manipulation is somewhat nebulous.
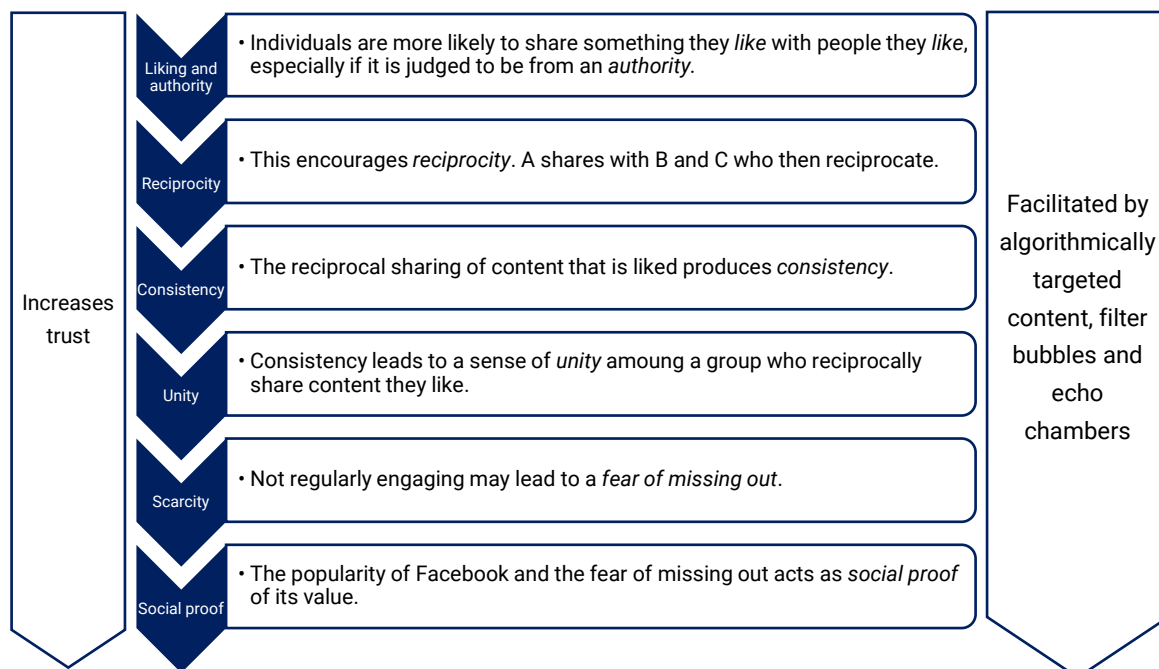

Figure 10: Cialdini's influence techniques applied to Facebook

---

[170] Robert B. Cialdini, *Influence: The psychology of persuasion* (New York: Harper Collins), 1984; Robert B. Cialdini, *Pre-suasion: A revolutionary way to influence and persuade* (New York: Random House), 2016.

Recall that sharing is integral to Facebook's idea of sociality and is a quantifiable means of increasing one's popularity. We are often motivated to respond in kind to pleasant or generous acts (returning the favour, so to speak) or, conversely, unpleasant acts. Reciprocity helps explain why some users are more likely to share information with others on Facebook who have previously shared their information with them (respond in kind).[171] Such 'friends' are people with whom we share *mutual obligations*.[172] Such sharing may also increase if the liked content is believed to have come from an authority. Not all users are motivated to engage in reciprocal sharing, however, or, if they do, it is not done for the sake of social cohesion. Robinson, for example, identified four types of Facebook user: *relationship builder*, *selfie*, *town-crier* and *window shopper*.[173] Of the four, only the 'relationship builder' was influenced to act based on reciprocity designed to increase or strengthen their social connections. Those categorised as 'selfies' shared information for narcissistic reasons: namely, as a means of self-presentation and to receive validatory responses. The 'town-crier' was primarily interested in posting information – opinions/comments – without seeking reciprocal exchanges. The 'window shopper' liked to collect information about others and was not interested in posting their own messages/comments (See also Appendix 6 for Moreno et al.'s Facebook influence model that presents four distinct categories of influence, some of which overlap with Robinson's findings).

Individuals are also more likely to share information they *like*, at least insofar as it appeals to them (i.e., it does not have to be a pleasant story or image), even if the information is suspected or even known to be false.[174] Liking something therefore influences how we respond to the information in terms of sharing it with others in our social network (who we like) which, in turn, is prone to elicit a reciprocal act of sharing from them, culminating in increased trust. The mechanism of reciprocity may also be a way of eliciting a response from a stranger, thereby increasing one's social capital on the platform,[175] and, given Facebook's set-up, one's popularity.

Facebook's business model encourages acts of reciprocal sharing. It is, after all, in Facebook's interest to increase prosumer productivity, a view echoed by van Dijck when he states:

> When communities of users continuously push "like" and "poke" buttons, issue recommendations, forward favorite items, or state preferences on what's hot and what's not, information generated in the informal connective sphere has "real" value on the commodity exchange floor.[176]

The reciprocal sharing of items of interest and the feedback the user receives, which in turn is shared with the rest of their social network, prompts Facebook to prioritise items that it anticipates will elicit similar sharing and feedback, the aim of which, as discussed, is to increase engagement and hence productivity.[177] The cumulative effect is to create a space that contains more and more targeted content. This, in turn, creates *consistency* (i.e., content that aligns with one's values and therefore one's 'likes' that others in one's network share). Consistency is a further means of influencing attitude and behaviour, given that people are prone to *confirmation bias* [178] – to seek

---

[171] Jong Gun Lee, Panayotis Antoniadis, and Kavé Salamatian, "Faving reciprocity in content sharing communities: A comparative analysis of Flickr and Twitter," Paper presented at the 2010 International Conference on Advances in Social Networks Analysis and Mining, Odense, Denmark, 9-11 August 2010; Jerzy Surma, "Social exchange in online social networks. The reciprocity phenomenon on Facebook," *Computer Communications* 73 (2016): 342-346.

[172] Robert D. Putnam, Bowling alone: The collapse and revival of American community (Simon and Schuster), 2000.

[173] Tom Robinson, Clark Callahan, Kristoff Boyle, Erica Rivera, and Janice K. Cho, "I ❤ FB: A Q-Methodology analysis of why people 'like' Facebook," *International Journal of Virtual Communities and Social Networking* 9, no. 2 (2017): 46-61.

[174] Jarred Prier, "Commanding the trend: Social media as information warfare," *Strategic Studies Quarterly* 11, no. 4 (2017): 50-85.

[175] Chris Chao Su, and Ngai Keung Chan, "Predicting social capital on Facebook: The implications of use intensity, perceived content desirability, and Facebook-enabled communication practices," *Computers in Human Behavior* 72 (2017): 259-268.

[176] van Dijck, "Facebook as a tool," 160-176. p.170.

[177] Facebook is able to employ sophisticated facial recognition software to help identify existing Facebook users or potential new users from photographs posted on its platform (see Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 2014, pp. 1701-1708, doi: 10.1109/CVPR.2014.220 for a detailed discussion on facial recognition software).

[178] Raymond S. Nickerson, "Confirmation bias: A ubiquitous phenomenon in many guises," *Review of General Psychology* 2, no. 2 (1998): 175-220; Peter C. Wason, "On the failure to eliminate hypotheses in a conceptual task," *Quarterly Journal of Experimental Psychology* 12, no. 3 (1960): 129-140.

out that which is taken to confirm their view – or adjust their attitude and/or behaviour to match those perceived to be similar.[179] Thus, it would seem that in order to be more readily exposed to a diverse selection of news items or other content, it is the *responsibility of the user* to diversify – in terms of other members' opinions and values – their social network.[180]

Fear of missing out (which falls under Cialdini's *scarcity* technique) can also influence attitudes and behaviour. In this case, users may continue to visit Facebook, or engage with the platform for longer periods, out of a fear of missing out on something important (e.g., the latest deal, or social event, or interesting news item).[181] Moreover, the likelihood that the user's Facebook network of 'friends' is, to some degree, a product of some or all of these influences means that Facebook acts like a filter.[182] One effect of this is to create a greater sense of *unity* (Cialdini's seventh, and most recent, persuasive technique) among those in one's network, or what Grimmelmann refers to as the influence of *community*.[183] As Grimmelmann notes, Facebook works hard to make its users feel that they are part of a trustworthy community of 'friends' and a feeling of familiarity with the other members is essential to this. A sense of unity or community can further influence one's attitudes and behaviour by creating a social bubble in which one receives affirmation from like-minded individuals (a further example of *consistency*) and avoids (regular) exposure to alternative views or content that challenges one's worldview, thereby adding to the filter effect that helped create the community in the first place and endorses one's *social identity*.[184]

Finally, the continuing success of Facebook – its ever-increasing popularity and revenue (recall Figure 2) – acts as a kind of *social proof*, for users old and new, that the platform is a good and therefore valued product worth joining and sticking with. Grimmelmann refers to this means of influence as *bigness*:[185] after all, "so many people can't be wrong!". To underscore this, Facebook is more than willing to promote its popularity: its bigness. This is because such social consensus implies trustworthiness. But, for Waldman, the 'like' button, as a means of collecting data points, crosses a divide between information we voluntarily give up – as part of our profile-building, image-management and means of increasing social capital – and the data Facebook collects by tracking us, and drawing inferences from our preferences, and so on.[186] To illustrate, Facebook algorithms currently work as follows:

— First, Facebook takes every post available in a user's network (a.k.a. the "inventory") and scores those posts according to predetermined ranking signals, like type of post, recency (etc.).

— Next, it discards posts that a user is unlikely to engage with, based on that user's past behaviour. It also demotes content that users do not want to see (e.g., clickbait, misinformation, or content that they have indicated they do not like).

[179] Rosanna, E. Guadagno, and Karen Guttieri, "Fake news and information warfare: An examination of the political and psychological processes from the digital sphere to the real world," In *Handbook of research on deception, fake news, and misinformation online*, edited by Innocent E. Chiluwa, and Sergei A. Samoilenko, 167-191. Hershey, PA.: IGI Global, 2019.

[180] Callum Borchers, "Facebook invites you to live in a bubble where you are always right," *The Washington Post*, 2018 Facebook invites you to live in a bubble where you are always right - The Washington Post.

[181] See Ine Beyens, Eline Frison, and Steven Eggermont, "'I don't want to miss a thing': Adolescents' fear of missing out and its relationship to adolescents' social needs, Facebook use, and Facebook related stress," *Computers in Human Behavior* 64 (2016): 1-8; Agata Błachnio, and Aneta Przepiórka, "Facebook intrusion, fear of missing out, narcissism, and life satisfaction: A cross-sectional study," *Psychiatry Research* 259 (2018): 514-19.

[182] Facebook uses algorithms to select information, and therefore *filter* what is made more readily available to the user and what is not. The act of filtering is based on a number of factors, such as whether the user has interacted with similar posts in the past (say, by liking, commenting or sharing them), how much other people have done the same and, importantly, whether the post is a paid advertisement targeted at the user. See Anja Bechmann, and Kristoffer L. Nielbo, "Are we exposed to the same "News" in the news feed?" *Digital Journalism* 6, no. 8 (2018): 990-1002; Philip Seargeant, and Caroline Tagg, "Social media and the future of open debate: A user-oriented approach to Facebook's filter bubble conundrum," *Discourse, Context & Media* 27 (2019): 41-48; Eli Pariser, *The filter bubble: What the internet is hiding from you* (Penguin UK), 2011.

[183] Grimmelmann, "Saving Facebook," 1137-1151.

[184] Henri Tajfel, "Social identity and intergroup behaviour," *Social Science Information* 13, no. 2 (1974): 65-93.

[185] Grimmelmann, "Saving Facebook," 1137-1151.

[186] Waldman, "Manipulating trust on Facebook," 175-198. See also Jason Koebler, and Joseph Cox, "Internal documents show Facebook has never deserved our trust or our data," *Vice*, https://www.usatoday.com/story/tech/columnist/komando/2016/03/04/dont-click-like-facebook-again-until-you-read/81264440/.

— Then it runs a more powerful neural network over the remaining posts to score them in a personalised way (e.g., Mona is 20% likely to watch tutorial videos from her chess group but 95% likely to post a heart reaction to a photo of her sister's new puppy) and ranks them in order of value.

— Finally, it arranges a cross-section of media types and sources so that a user has an interesting variety of posts to scroll through.[187]

In summarising the discussion thus far, Luhmann famously remarked that trust exists where knowledge ends.[188] Somewhat ironically, perhaps, Facebook has devised a system of knowledge acquisition, or at least a means of collecting data points that can be converted into useful and marketable information, arguably by *manipulating* our willingness to trust, or at least influencing it through techniques of non-rational persuasion. This is because Facebook does not simply *utilise* the same mechanisms of influence that would no doubt facilitate an increase in social capital, trust and popularity, even if performed in more traditional social spaces; rather, it seeks (arguably) to exploit them and those who are influenced by them for the sake of increased commodity. Facebook's social space, and the trust it depends on, has therefore not evolved organically but has, instead, been shaped, and continues to be shaped, by its business model[189] and the psychological techniques of influence it cultivates to create its own brand of social engagement in an environment that, at the very least, favours the use of decision-making heuristics and other forms of non-rational persuasion. These techniques have been adopted and refined, one might surmise, because they enable users to act as prosumers, engaging in the kinds of digital labour that count as the means of production for Facebook, and the extent to which a technique is successful is determined by the extent to which it promotes prosumerism.

Whether the use of these techniques has always been deliberate, however, at least in the early days of Facebook – based on the intentional application of psychological theory – is difficult to gauge. The fact that Facebook has commissioned research looking into factors that influence engagement does suggest intent on the part of the organisation – at least *now* – even *if,* previously, the platform's compatibility with these techniques was unintended and merely fortuitous (for Facebook). It is worth noting, however, that BJ Fogg, who pioneered much of the research on mass influence techniques in relation to computers – during a period that predates but includes the emergence of Facebook – spent time, in addition to lecturing on the subject at Stanford University, consulting with tech industry. He has thus become something of a guru to those in Silicon Valley.[190]

How does Facebook enable third-party customers to influence their target audience?

What most, if not all, users of Facebook know is that they provide the company with all sorts of information, known as *transparent personalisation*.[191] What most are unaware of, however, is just how much Facebook is able to infer (through big data and pattern analysis) from all of the billions of small data points it continuously accumulates about individual preferences and tendencies, and therefore what someone or some group's likely attitude will be towards *x*, and even their subsequent behaviour. Individual data points are used to inform big data and algorithms that seek to influence users' attitudes and behaviour, and because this data-derived knowledge is of great use to advertisers and other third-parties it is a valued asset.[192]

As already discussed, Facebook users are not an audience in any typical sense; rather, they serve as the means by which Facebook acquires its valued commodity. Neither are they *customers*, given that Facebook membership and platform use is free, at least in terms of a lack of monetary requirement. Of course, users incur a cost in other ways (i.e., by handing over information beyond transparent personalisation, often unwittingly, in the form of personal data

---

[187] Points adapted from https://blog.hootsuite.com/facebook-algorithm/.
[188] Niklas Luhmann, *Trust and Power* (New York: Wiley), 1979.
[189] See Roel Wieringa, and Jaap Gordijn, "Facebook business models," TVE: The Value Engineers. 2020 Team - The Value Engineers for a detailed overview.
[190] Psychologists speak out against tech companies that use "persuasive design" to lure in kids - Vox.
[191] Balfour, "Likes, comments, action!" 34-51.
[192] Koebler, and Cox, "Internal documents show Facebook has never deserved our trust or our data."

points). Facebook's customers are, instead, those companies and organisations willing to pay for what Facebook's data and artificial intelligence gives them access to: namely, a target audience.[193]

Owing to Facebook's extensive knowledge of its users' habits, inclinations, tastes and so on, advertisements have tremendous targeting potential. Advertisements can be (and often are) personalised, so as to be perceived as non-intrusive purchasing advice from a 'friend' who, users are led to believe, also has an interest in the product (see Appendix 7).[194] Indeed, Facebook users may have difficulty distinguishing between a friend's post and a sponsored advertisement presented so that it appears as if it is content that has been shared by a friend,[195] or even a celebrity.[196] As Waldman explains:

> Native advertisements, or third-party links that are designed to look like social posts, also appear on our Newsfeeds. Like the social posts of our friends, these ads are often preceded by the names of our friends who have "liked" the advertiser's page… The information about our friends, not the ad, is the first thing we see. The only thing that distinguishes these ads from our friends' social posts is the word "Sponsored," written in light grey text under the name of the company and sandwiched between the ad's much larger graphic content and Facebook's note about our friends.[197]

Elsewhere, Waldman makes a related point:

> Privileging the posts of friends and family over the posts of third-party publishers… may limit the reach of a naked post from [these third-party publishers (i.e., a particular company, advertising their product)], but not when one of their videos is shared by a friend. That is ideal for Facebook for two reasons. First, users tend to dislike seeing posts from third-parties; second, under the new design, most third-party content that users see will come through their trusted social networks of friends. This cues the trustworthiness of the post far better than any naked post from a publisher ever could.[198]

Such digital manipulation can be individually tailored to appeal to users based on Facebook's pervasive data collections (one can therefore target those most susceptible to manipulation). Anspach, for example, found that a friend's (seeming) endorsement of content – a news item, say, or an advertisement – acted as a heuristic when deciding whether or not to read it.[199] This is because interpersonal communication from *opinion leaders* (recall Cialdini's *authority* technique for persuasion) may help *opinion followers* make sense of and evaluate the news content or other items shared on Facebook.[200] Moreover, in a series of field studies, Matz et al. found that psychologically tailored advertising – that is, advertising presented in a way geared to appeal to an individual based on their psychological make-up (e.g., an extrovert) – produced significantly more clicks and purchases than advertising presented in a manner inconsistent with an individual's psychological make-up.[201] Kruikemeir et al., for

---

[193] White and Boatwright, "Social media ethics," 1-7. See also Elizabeth Van Couvering, "Faces and charts: platform strategies for visualising the audience, the case of Facebook, Information," *Communication & Society*, (2021): 1-18. for an analysis and discussion on the differences in the way Facebook interfaces with users (or members), publishers and advertisers, each of which produces difference data.

[194] Fumagalli et al., "Digital labour," 1757, 1-16.

[195] Shrum, L.J., The psychology of entertainment media: Blurring the lines between entertainment and persuasion (New York: Routledge), 2012.

[196] Sophie C. Boerman, Lotte M. Willemsen, and Eva P. Van Der Aa, "'This post is sponsored': Effects of sponsorship disclosure on persuasion knowledge and electronic word of mouth in the context of Facebook," *Journal of Interactive Marketing* 38 (2017): 82-92.

[197] Waldman, "Manipulating trust on Facebook," 175-198, p.194.

[198] Waldman, "Privacy, sharing, and trust," 193-234. p.223.

[199] Anspach, "The new personal influence: How our Facebook friends influence the news we read," 590-606. See also Veronika Karnowski, Anna S. Kümpel, Larissa Leonhard, and Dominik J. Leiner, "From incidental news exposure to news engagement. How perceptions of the news post and news usage patterns influence engagement with news articles encountered on Facebook," *Computers in Human Behavior* 76 (2017): 42-50.

[200] Jason Turcotte, Chance York, Jacob Irving, Rosanne M Scholl, and Raymond J Pingree, "News recommendations from social media opinion leaders: Effects on media trust and information seeking," *Journal of Computer-Mediated Communication* 20, no. 5 (2015): 520-535.

[201] Sandra C. Matz, Michal Kosinski, Gideon Nave, and David J Stillwell, "Psychological targeting as an effective approach to digital mass persuasion," *Proceedings of the national academy of sciences* 114, no. 48 (2017): 12714-19.

their part, reported that posts not noticed to have been sponsored by a political party were treated the same as posts by a friend and shared.[202] Where the sponsorship was noticed, however, the content tended not to be passed on to friends.

Facebook uses an immune system algorithm to control its users' "mediated experience towards a desired rhythm (sociality) while filtering out problematic rhythms (spam)",[203] (see Figure 11). These rhythms are a marketable product for Facebook because they are of value to advertisers who bid for the data so that they can intervene to shape people's experience at the most opportune times (that is, in a manner that accords with these rhythms). Facebook therefore shapes, manages and filters specific rhythms as a means of ordering sociality to make it more valuable.[204] Thus:

> With the Facebook Immune System algorithm, Facebook listens to people's behaviors, whether they are considered positive or negative – everything counts. In this way, the company establishes what types of rhythms can harm its business and thus should be decreased/removed/filtered as possible options of behaving on the platform. The platform only orchestrates people in the desired rhythm, so although all actions count, only the valuable behaviors (such as more emotional or repetitive) will be ordered whilst the 'negative' ones (such as creepers) will be filtered out.[205] In this way, the company establishes what counts as engagement and what type of sociality is worth more.[206]

These data enable advertisers to target specific audiences with products and services Facebook 'knows' (based on inferences made from our online behaviour) that we want or are at least interested in. Facebook therefore not only helps the advertiser identify a potential audience but seeks to ensure that, once identified, members of the audience are targeted by the advertisement in a personalised way.[207]

Facebook's marketing service "Lookalike Audiences," for example, goes beyond targeting users' based merely on their 'likes' and preferences. Instead, it categorises them based on other factors or characteristics of interest to a particular third party, such as spending capacity. To illustrate: Muka, a casino game developer, used the service to target "high-value players", characterised as those "most likely to make in-app purchases" (according to Singer, reporting for *The New York Times*).[208]

---

[202] Sanne Kruikemeier, Minem Sezgin, and Sophie C Boerman, "Political microtargeting: Relationship between personalized advertising on Facebook and voters' responses," *Cyberpsychology, Behavior, and Social Networking* 19, no. 6 (2016): 367-372.
[203] Heejin Lim, and David W. Schumann, "Employing a Dramaturgical Lens to the Interpretation of Brand Online Social Networking: Evidence of Augmented Self," *Qualitative Market Research: An International Journal* 22, no. 3 (2019): 278-300. See also Tao Stein, et al., "Facebook immune system," *EuroSys Social Network Systems (SNS)*, Salzburg, 10 April 2011: 1-8. facebook-immune.pdf (mit.edu).
[204] Lim and Schumann, "Employing a Dramaturgical Lens to the Interpretation of Brand Online Social Networking: Evidence of Augmented Self," 278-300.
[205] 'Creepers' is a name given to 'normal' Facebook users who do not use the platform according to its 'intended uses' (as determined by Facebook). From Facebook's perspective, then, creepers harm the company's business model.
[206] Lim, and Schumann, "Employing a Dramaturgical Lens to the Interpretation of Brand Online Social Networking: Evidence of Augmented Self," 278-300.
[207] See https://adespresso.com/guides/facebook-ads-beginner/demographic-targeting/ and https://www.wordstream.com/blog/ws/2017/11/01/facebook-audiences.
[208] Natasha Singer, "What you don't know about how Facebook uses your data," *The New York Times*, 11 April 2019. https://www.nytimes.com/2018/04/11/technology/facebook-privacy-hearings.html.
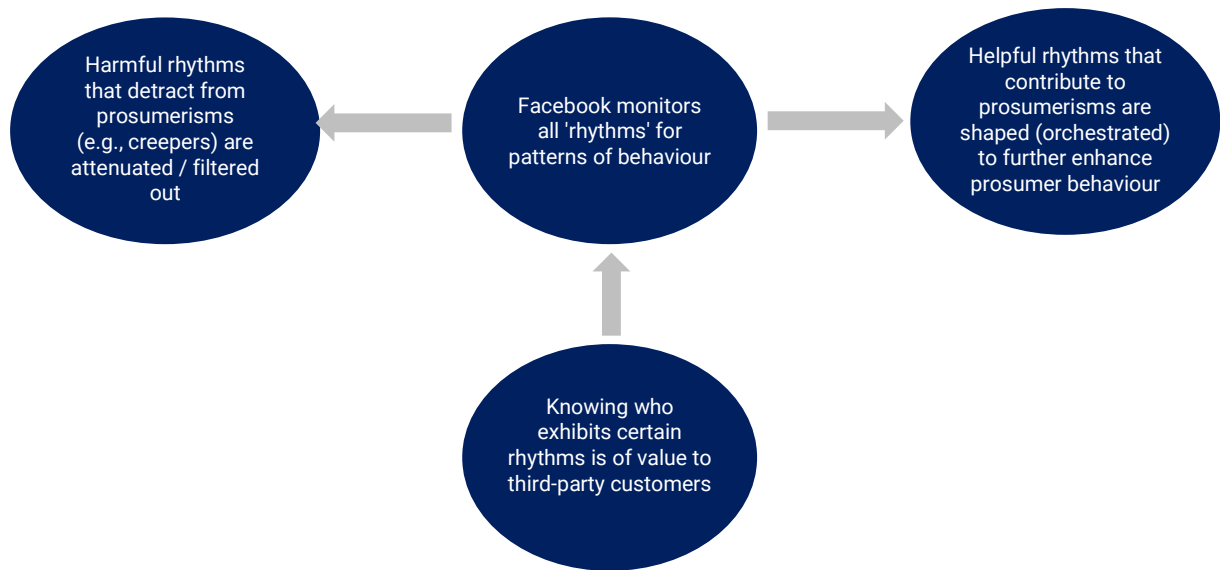
Figure 11: Monitoring and shaping users' patterns of behaviour.
Facebook monitors users' patterns of behaviour or 'rhythms' and seeks to shape these.
Access to users (target audiences) that exhibit certain rhythms is a value to third-party organisations.

Facebook also provides direct access to the audience members' pages.[209] Fan pages – namely, websites within Facebook that are created by marketers – are also an effective means of personalising content and enhancing engagement.[210] Users become 'fans' when they link their personal profiles to a fan page. Fan pages contain a 'wall' where members and a particular company (e.g., Audi) can post content. Marketers regularly post content soliciting feedback to engage fans (e.g., "What currently unavailable colour would you like to see made available?")

Tailored influence (or even manipulation) will continue to improve as ever advancing analytics show which type of intervention is most effective over time. This means that given advances in digital technology, Facebook will be able to adapt continually to its prosumers' responses and have the potential to make 'targeted influence' an ongoing process, none of which need be transparent to prosumers.[211]

In summarising this section: Advertising on Facebook seems to be more and more dependent on targeting the 'audience' that is most amenable to what the third-party customer is selling. "Facebook uses data collected from users to ensure that the right advertisement will reach the right user at the right time. Every time a user visits Facebook, an algorithm will choose between thousands of ads and will show those that it deems the most relevant… to each user profile."[212] The peers and close online connections of the targeted audience are also seen as part of the advertising 'audience', and these social connections of trust and influence are vulnerable to non-rational persuasion techniques and manipulation. When requests for responses are interesting and fun, and give the 'fan' the opportunity for self-expression, then they tend to create a positive impression of the marketed product.[213] Interactions can also "make individuals feel connected with an interaction partner (or a fan page) and may increase perceived familiarity, which also increases liking".[214] Fan pages help Facebook promote a consumer-to-consumer approach, whereby users are encouraged to share experiences, and create a common pool of knowledge on products and services; but, equally, they provide businesses and brands with a direct means of communication to

[209] Bernal, "Facebook," 513-530.
[210] Philipp A. Rauschnabel, Sandra Praxmarer, and Bjorn S Ivens, "Social media marketing: How design features influence interactions with brand postings on Facebook," In *Advances in Advertising Research (Vol. III)*, 153-161: Springer, 2012.
[211] Tal Z. Zarsky, "Privacy and manipulation in the digital age," *Theoretical Inquiries in Law* 20, no. 1 (2019): 157-188.
[212] Cetina Presuel, Rodrigo, and José Martínez Sierra, "Algorithms and the News: Social Media Platforms as News Publishers and Distributors," *Revista De Comunicación* 18, no. 2 (2019): 261-285, p.276.
[213] Kathy N. Shen, and Mohamed Khalifa, "Exploring Multidimensional Conceptualization of Social Presence in the Context of Online Communities," *International Journal of Human-Computer Interaction* 24, no. 7 (2008): 722-748.
[214] Rauschnabel et al., "Social media marketing," 153-161. p.156.

an already interested audience, thereby enabling Facebook to adopt a business-to-consumer approach as well.[215] Brand Online Social Networking (BOSN), for example, is a novel means of marketing that enables companies to initiate and cultivate relationships with their customers through social media platforms like Facebook.[216] Success in the *market for attention* (which is motivated by the expectation that attention will lead to further engagement) is therefore increasingly the result of personalised – targeted – algorithmic selection procedures, the net effect of which is audience fragmentation, and the creation and enhancement of individuated or homogenous group realities.[217]

---

[215] Loredana Di Pietro, and Eleonora Pantano, "An empirical investigation of social network," Journal of Direct Data and Digital Marketing Practice 14 (2012): 18-29.

[216] Lim, and Schumann, "Employing a Dramaturgical Lens to the Interpretation of Brand Online Social Networking: Evidence of Augmented Self," 278-300.

[217] Natascha Just, and Michael Latzer, "Governance by algorithms: reality construction by algorithmic selection on the Internet," *Media, Culture & Society*, 39, no. 2 (2017): 238-258.

Theme 3: Systems and Technology

Research Question 3: What are Facebook's foundational systems, and what technology and workforce skills are required for its operation now and in the future?

Facebook is influential not only because it successfully employs a very effective business model based on the collection of data, the creation of profiles and the sale of targeted ads, but also because it is an entity whose influence extends across platforms.

Facebook's foundational systems and technology

Facebook uses machine learning to predict which content will matter most to each person and ranks its importance according to the user's responses (such as 'likes' and shares) to other content (such as friends' posts) so that users will find the content more engaging and have positive experiences. To do this, multiple prediction models for each user, based on different information related to the user, are created. These predictions combine to form a single ranking score for each post with respect to the user. Most importantly, Facebook does this very efficiently, in real time, on a system infrastructure that can scale to 2+ billion people with more than 1000 posts per user per day.

Face recognition

Designed in-house, *DeepFace* is a facial recognition tool with a higher accuracy rate than humans – 97% compared with 96% – and the FBI's Next Generation Identification system, which is correct 85% of the time. It is used to analyse the photos and videos an individual is in that are posted on Facebook, such as their profile picture, and photos and videos that they have been tagged in. It does this to make a unique user number called a template. Facebook introduced its face recognition in 2015, except in the EU, where data privacy legislation prevents it from being used.[218]

Augmented reality ads

Augmented reality (AR) is the placement of computer-generated imagery in a user's field of view. Facebook AR Ads are advertisements in a user's newsfeed, targeted at them and their interests. With AR technology, one can even "try it on," or at least see what it looks like on one's face or body via the ads. It is based on Facebook's Spark AR,[219] a studio tool that allows users to create their own AR effects. Launched in 2017, Facebook continues to add capabilities (e.g., analytics) to the platform. Spark AR allows the user to create their own AR effects using a suite of tools, from patching to animation, in a user-friendly manner. It enables a person to create anything from a face filter to interactive AR games. Together with the analytics capabilities, Spark AR provides insights into how the user responds to AR.[220]

The range of technologies that will be required by the Australian Defence Force (ADF) will clearly require a considerable amount of future research. This will have to be done in conjunction with a determination of the specific functions of such a centre and the appropriate parameters that should contain it.[221]

---

[218] Yaniv Taigman; Ming Yang; Marc'Aurelio Ranzato; Lior Wolf, DeepFace: Closing the Gap to Human-Level Performance in Face Verification, IEEE Conference on Computer Vision and Pattern Recognition, 2014. https://ieeexplore.ieee.org/document/6909616.

[219] https://sparkar.facebook.com/ar-studio/.

[220] Diago Aguirre, "Ten Things You Need to Know About Spark AR," *Diego Aguirre*, 13 May 2019.

[221] Daniel Carnahan, "Facebook has made AR ads available to all marketers through its ad manager," *Insider*, 13 December 2019.

The Facebook workforce

Facebook is unwilling to share information on its technology or workforce as it regards it as commercial intelligence. It prevents the spread of intellectual property by siloing its workforce. Researchers looking into Facebook's workforce have had only limited success, and determining Facebook foundational systems, technology and workforce skills is difficult.[222] In addition, we need to be cautious about describing Facebook within a particular era as it is the organisation's power to *adapt* that has helped maintain its dominance. It is, therefore, important to understand the enduring characteristics that have made Facebook's operations successful under several different contexts. That said, Facebook shares some recruitment practices and there are many Facebook-based references on what makes a good workforce. Zuckerberg has been quoted as saying that the "single most important thing" is to get the best people, contending that CEOs need to set aside their egos and hire people they admire.[223]

In December, 2020 Facebook claimed a global workforce of 58,604[224] excluding those working through contract arrangements. Those employed directly generally enjoy high-quality conditions, including good pay, long vacation leave, health and dental care, good paternity leave, a "wellness allowance" for gym membership and a "baby cash" payment for new parents. Other perks include free meals, a dry-cleaning service, valet parking, a video games arcade and a bike repair shop.[225]

Employees are clearly expected to display a positive attitude towards the company and their positions in it. In a number of company-approved articles there is a great deal of emphasis placed on staff development and wellbeing, as well as a shared vision. One employee cites Facebook's "mission, culture, perks, and Zuckerberg" as the reason he enjoys working at Facebook. Others cite the fact that the company "gives people the ability to focus on what they love", utilises a "strengths-based management philosophy" and "values passionate people regardless of expertise". Some cite Facebook's "core mission of connecting the world" as a motivating factor, while another emphasises the "really cool people" he met during his employment interview.[226] A data scientist describes feeling like "the luckiest guy on earth for landing a job here" and a software engineer claims that "Facebook has truly surpassed all tech companies in terms of culture, perks, and employee lifestyle".[227]

A number of current and former Facebook employees speaking anonymously, however, provide a different account.[228] They contend that Facebook's culture discourages dissent, holds back employees who lack enthusiasm for the company and punishes those who leave. One complained of pressure to attend team-building events outside of work hours, even under difficult circumstances. Some complain of the employee evaluation system, whereby employees must choose five colleagues to evaluate them, producing an atmosphere of conformity. Facebook also uses the controversial "stack ranking" system. The performance of employees is sorted into seven categories with defined percentages in each. Those in the lowest 15% are in danger of being dismissed. Employees who leave or are dismissed are classed as "regrettable" or "not regrettable." The latter are not re-employed and are provided with bad references. Zetlin sums up Facebook's employment culture as:

— You must keep Facebook secrets

— You must be available outside work hours

— You should not openly disagree with your boss or your peers

— You must express enthusiasm and love for the organisation both internally and publicly

[222] Discussion between the UNSW Canberra research team and Lise Waldek and Julian Droogan, 4 March 2021.
[223] Richard Feloni, "Mark Zuckerberg shares the hiring rule he says separates good companies from great ones," *Business Insider*, 13 June 2017. https://www.businessinsider.com.au/mark-zuckerberg-hiring-rule-2017-6?r=US&IR=T.
[224] "Form 10-K. Annual report pursuant to section 13 or 15(D) of the Securities Exchange Act of 1934. For the fiscal year ended 31 December 2020. Facebook, Inc," (Washington: United States Securities and Exchange Commission, 2020), 11.
[225] Áine Cain, "The 11 coolest perks at Facebook, the best place to work in America," *Business Insider*, 7 December 2017. https://www.businessinsider.com.au/facebook-employee-perks-benefits-2017-11?r=US&IR=T#a-bike-repair-shop-for-menlo-park-employees-2.
[226] Áine Cain, "What it is REALLY like to work at Facebook," *Business Insider*, 7 December 2017. https://www.businessinsider.com.au/what-its-really-like-to-work-at-facebook-2017-12?r=US&IR=T.
[227] Rachel Gillett, "7 reasons Facebook is the best place to work in America," *Business Insider*, 7 December 2017. https://www.businessinsider.com.au/facebook-best-place-to-work-in-america-2017-12?r=US&IR=T.
[228] Facebook culture described as 'cult-like', review process blamed (cnbc.com).

— You are in constant danger of being cast out and if you are cast out you have a hard time finding a place elsewhere

— You must be very careful about how you leave[229]

The benefits provided to official Facebook employees are not extended to those under subcontracting arrangements, with those at the Menlo Park campus receiving a basic minimum of US$15 an hour. In a region of high costs, many are forced to take second or third jobs and some are homeless.[230] The international subcontracted team of Facebook moderators, referred to by Facebook as "process executives" and numbering some 16,000 in 2019, undertake the traumatising task of moderating what can be very disturbing material. While, for years, moderation occurred in countries such as the Philippines due to wage differentials, in more recent years Facebook has utilised more subcontractors in the US, where they are considered more sensitive to US cultural values. In Phoenix, for example, the subcontractor Cognizant employs around 1000 moderators, mostly on a casual basic wage of US$15 an hour, averaging US$28,800 a year. They are required to sign nondisclosure agreements to protect Cognizant and Facebook from criticism about their working conditions. Employees speaking confidentially to *Verge* journalist Casey Newton described an environment "teetering on the brink of chaos", in which employees can be easily dismissed, bathroom and prayer breaks are micromanaged, and private phones confiscated during work hours. Newton concludes that due to the traumatising nature of the material they work with, many struggle with trauma long after they have left their jobs and what counselling is provided ends when they leave.[231]

Psychologists maintain that viewing violent, pornographic and child abuse material can be significantly detrimental to moderators' well-being.[232] A former Facebook analyst contends that support services are grossly inadequate and that workers need resiliency training and counselling. Facebook claims to offer psychological support but there is no mandatory counselling. One employee claims many feared they would lose their jobs if they sought psychological help through the company. Solon contends that Facebook practices fall short of industry standards, citing the example of the UK Internet Watch Foundation (IWF), which has detailed training and support for such employees, including psychological evaluation before hiring and extensive resilience training.[233]

How has Facebook adapted its communication strategies to maintain dominance?

Tillman lists 10 reasons Facebook has achieved and maintained its global dominance (see Table 2):[234]

Table 2: Ten reasons for Facebook's continued dominance

1.	Ease of use	Facebook is relatively easy to use, even for those who feel less competent with digital technology.
2.	Constant upgrades	Facebook has always been willing to develop upgrades. At the F8 2019 conference, for example, Facebook executives announced plans to release a redesign called *New Facebook,* an interface that will place a greater emphasis on Groups and Events: two of the biggest reasons people visit Facebook.

---

[229] Minda Zetlin, "Here's Why Facebook's Former Employees Describe the Company as Cult-Like," *Inc*, Undated. https://www.inc.com/minda-zetlin/facebook-culture-cult-like-former-employees-report.html.
[230] Julia Carrie Wong, "Facebook's underclass: as staffers enjoying lavish perks, contractors barely get by," *The Guardian*, 26 September 2017. https://www.theguardian.com/technology/2017/sep/26/facebook-workers-housing-janitors-unique-parsha.
[231] Casey Newton, "The Trauma Floor: The secret lives of Facebook moderators in America," *The Verge*, 25 February 2019. https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona.
[232] Olivia Solon, "Facebook is hiring moderators. But is the job too gruesome to handle?" *The Guardian*, 4 May 2017. https://www.theguardian.com/technology/2017/may/04/facebook-content-moderators-ptsd-psychological-dangers.
[233] Olivia Solon, "Underpaid and overburdened: the life of a Facebook moderator," *The Guardian*, 25 May 2017. https://www.theguardian.com/news/2017/may/25/facebook-moderator-underpaid-overburdened-extreme-content.
[234] Tillman, Maggie, "10 reasons why Facebook has thrived for 15 years," Pocket-lint, 4 February 2019. 10 reasons why Facebook has thrived for 15 years (pocket-lint.com).

3.	New features	Facebook regularly invests in new features (such as Spark AR). Some are more successful than others. [235]
4.	Compatibility with mobiles	Facebook quickly adapted to be compatible with smartphones.
5.	Successful acquisitions	Facebook has sought to assert dominance by acquiring successful companies (see Appendix 2), the most notable being WhatsApp, Instagram and Oculus. Facebook has been accused of adopting a very aggressive strategy at times. [236]
6.	Worldwide and generational users	Facebook is a worldwide phenomenon. It also has cross-generational appeal.
7.	Seeking to eliminate anonymity	By requiring an authentic account, Facebook has sought to eliminate anonymity. This gives the platform the appearance of being safe and inviting.
8.	Facebook Connect	*Facebook Connect* allows Facebook users to use their account across different platforms, rather than having to set up new accounts each time. This is a means by which Facebook ingratiates itself with users and other systems.
9.	Becoming a public company	Becoming a public company enabled Facebook to increase investment and compete with companies such as Apple, Google, Amazon and Microsoft.
10.	Adapting to accommodate video	Facebook has adapted to accommodate the rise in popularity of video, initiated by YouTube.

In addition, Doody notes that Facebook has a good public relations team,[237] not only when faced with market scrutiny and outcries such as the Cambridge Analytica revelations but also when confronted with evidence that extensive social media use is correlated with declining mental health.[238] In response to the mental health issue, Facebook has acknowledge evidence of risk but also promoted the counter-claim that Facebook engagement can have a positive mental health impact.[239] It has also devised and promoted its suicide-post detection tool, as well as Facebook mental health support groups.[240]

Doody also notes that Facebook has been able to maintain its dominance because other large tech-companies – such as Amazon, Google and LinkedIn – do not mimic Facebook's approach and do not engage in direct competition. When Google tried to compete directly with Facebook (via its 2011-2019 ill-fated Google+ platform) it failed.[241] Nevertheless, for as long as each of these companies continues to maintain their own advertising capabilities, they will pose a potential threat to Facebook's main source of income. WeChat – by virtue of being the largest social network in China with over 1.15 billion monthly users – likewise has the potential to threaten Facebook's dominance.[242] There will also always be things that Facebook cannot control, such as government

---

[235] Facebook's messaging platform is very popular. However, Facebook Beacon was not, eventually being removed after a lawsuit. Facebook Beacon reported (to Facebook) users' activity in third-party sites. It did this without gaining the user's permission.

[236] Facebook undermined rivals in bid to dominate global messaging (computerweekly.com).

[237] See also Gil Press. "Why Facebook triumphed over all other social networks," Forbes, 8 April 2018. Why Facebook Triumphed Over All Other Social Networks (forbes.com).

[238] Hui-Tzu (Grace) Chou, and Nicholas Edge, "'They are happier and having better lives than I am': The impact of using Facebook on perceptions of others' lives," *Cyberpsychology, Behavior & Social Networking* 15, no. 2 (2012): 117-121; Ethan Kross, et al., "Facebook use predicts declines in subjective well-being in young adults," PLoS ONE 8 no. 8 (2013): e69841. Facebook Use Predicts Declines in Subjective Well-Being in Young Adults (plos.org).

[239] For useful counter findings and reviews on research, see Kathleen Kjelsaas, et al., "For better or worse: Facebook use and wellbeing among community health care clients," The Journal of Nervous and Mental Disease 2007, n. 8 (2019): 659-667; Igor Pantic, "Online Social Networking and Mental Health," Cyberpsychology, Behavior, and Social Networking 17, No. 10 (2014): 652-657.

[240] How Facebook AI Helps Suicide Prevention - About Facebook (fb.com); 7 Facebook Pages to Follow about Mental Health Support and Education - Salience (salienceneuro.com).

[241] Google begins shutting down its failed Google+ social network - The Verge; Why Google+ can never compete with Facebook | Network World.

[242] 6 Powerful WeChat Statistics in 2020 You Need to Know (brewinteractive.com).The following papers provide informative discussion on similarities and differences between Facebook and WeChat, particularly in relation to user experience: Hanjing Huang, et al., "International Users' Experience of Social Media: A Comparison Between Facebook and WeChat," **10th**

regulation. In 2019, for example, the FTC inquired into whether Facebook defensively purchased certain companies (Instagram and WhatsApp) in order to maintain its dominant market position.[243] In 2020, the FTC sued Facebook for what it claimed were illegal monopolisation practices.[244]

The reason Facebook continues to dominate in the arena of social networking is described by Doody:

> For advertisers and brands, [Facebook] is used as an outlet to communicate, build communities, drive sales, shape opinions and ultimately grow businesses, making it an integral part of any marketing strategy…, often outperforming more predictable means of advertising.[245]

Government Liaison with Facebook

Examples of cooperative activity between social media entities and government authorities exist. One is the Global Internet Forum to Counter Terrorism (GIFCT), founded by Facebook, Google, Microsoft and Twitter in August, 2017.[246] The forum was established to foster cooperation between companies, advance relevant research and engage with other stakeholders, including governments, to counter the spread of terrorism, and extremist and violent content, online.[247] Another example is the *Christchurch Call to Action*, formed by governments after the March, 2019 mosque shootings in Christchurch, New Zealand. Tech companies have signed onto the initiative's plan to coordinate industry efforts to combat violence and extremist material online.[248]

To facilitate an effective liaison facility, it is necessary to consider how law enforcement liaison operates currently. The Telecommunications Act 1997 (Cth) (Telco Act) defines three types of service provider in the telecommunications space. The first is a carriage service provider delivering communications services to the public. The second is a carrier, which is a carriage service provider that delivers communications services using its own infrastructure. The key point is that carriers own infrastructure, but a carriage service provider might simply resell services. For example, ALDImobile is a carriage service provider which uses the Telstra network and Telstra is a carrier. The third type of service provider is a content service provider delivering content services to the public using a carriage service. A content service is broadly defined and includes Internet content such as Facebook pages.

Under the Telco Act, anyone who operates a "network unit" to supply a carriage service to the public must hold a carrier licence. As network units are the basic units of telecommunications infrastructure, this is how carriers and carriage service providers are separated. However, there is an issue. Carriers hold licences, and the Australian Communications and Media Authority (ACMA) is required to maintain a register of licensed carriers.[249] Carriage service providers and content service providers, on the other hand, do not have to be registered or licensed, although they have obligations that flow from the class of services they provide. This concept of "class licence" and "class licensing" is used in the *Radiocommunications Act 1992* (Cth) (Radcomms Act) and the *Broadcasting Services Act 1992* (Cth) (BSA).

International Conference, CCD 2018. Held as Part of HCI International 2018 Las Vegas, NV, USA, 15–20 July 2018, Proceedings, Part I, Pei-Luen Patric Rau (ed.): 341-349; Chunhui Xie, et al., "Family, Friends, and Cultural Connectedness: A Comparison Between WeChat and Facebook User Motivation, Experience and NPS Among Chinese People Living Overseas," 9th International Conference, CCD 2017 Held as Part of HCI International 2017 Vancouver, BC, Canada, 9–14 July 2017, Proceedings. 369-384.

[243] Katie Paul, and Jan Wolfe, "U.S. FTC probes Facebook's acquisition practices," *Reuters*, 2 August 2019. U.S. FTC probes Facebook's acquisition practices - WSJ | Reuters.

[244] Federal Trade Commission, (US), "FTC Sues Facebook for Illegal Monopolisation," *Federal Trade Commission press releases*, 9 December 2020.

[245] Amy Doody, "Can anyone challenge Facebook's dominance in social?" *The Irish Times*, 7 February 2020, p.1 Can anyone challenge Facebook's dominance in social? (irishtimes.com).

[246] Global Internet Forum to Counter Terrorism. https://gifct.org/about/.

[247] David Lee, "Tech forms hail 'progress' on blocking terror", BBC News, 8 June 2018. https://www.bbc.com/news/technology-44408463.

[248] *Christchurch Call to eliminate terrorist & violent extremist content online*, New Zealand Ministry of Foreign Affairs and Trade. https://www.christchurchcall.com/call.html.

[249] This is at https://www.acma.gov.au/register-licensed-carriers.

Operating a radio transmitter in Australia without a licence is prohibited under the Radcomms Act and s.51(v) of the Constitution, which reserves powers over "postal, telegraphic, telephonic, and other like services" to the Commonwealth. However, it would be impractical to individually license every WiFi router. Instead, WiFi routers that meet the technical specifications set out in the *Radiocommunications (Low Interference Potential Devices) Class Licence 2015* fall into a class of device that is permitted to be used without an individual licence. Similarly, under the BSA, ACMA has the power to create class licences for subscription narrowcasting services, to avoid the regulatory burden of having to apply for a subscription broadcasting services licence.

The reason for labouring this point is that Facebook already has obligations under the Telco Act as a content service provider which are binding whether Facebook is aware of this or not. Obligations on carriage service providers and content service providers, called "service provider rules", are set out in section 98 of the Act. Those on carriage service providers are not relevant to this discussion. However, in regard to content service providers, the following obligations apply:

a)   the rules set out in Schedule 2 of the Telco Act; and

b)   the rules (if any) set out in service provider determinations in force under section 99 of the Telco Act.

The relevant rule in Schedule 2 is "comply with the Telco Act". There is also an obligation to comply with Chapter 5 of the *Telecommunications (Interception and Access) Act 1979* (Cth). In practice, this only applies to carriage service providers.

Under section 99 of the Telco Act, ACMA may make a service provider determination in respect of a particular type of content service. In doing so, it must have specific regard to the Constitution and it must consult with the Australian Competition and Consumer Commission (ACCC). It would be feasible for ACMA to use a definition developed by the ACCC in the Digital Platforms Inquiry[250] to specify Facebook's content service and have a service provider rule associated with that content service.

Carriage service providers have an obligation which is currently used to create law enforcement liaison units in major carriers and carriage service providers. This flows from section 313 of the Telco Act. Section 313(3) relevantly provides:

A carrier or carriage service provider must, in connection with:

a)   the operation by the carrier or provider of telecommunications networks or facilities; or

b)   the supply by the carrier or provider of carriage services;

give officers and authorities of the Commonwealth and of the States and Territories such help as is reasonably necessary for the following purposes:

a)   enforcing the criminal law and laws imposing pecuniary penalties;

b)   protecting the public revenue;

c)   safeguarding national security.

Section 313(3) of the Telco Act was examined by the House of Representatives Standing Committee on Infrastructure and Communications in the context of online harms.[251] However, the review did not consider whether the obligations of section 313(3) of the Telco Act should be extended to content service providers. To some extent, this was a result of the terms of reference of the inquiry.

What we can learn from Facebook

The Russian campaign of disinformation during the 2016 US presidential elections worked across platforms, incorporating mediums such as Twitter, but also broader contexts such as talkback radio. While seeding content

---

[250] ACCC, "Digital Platforms Inquiry - Final Report" (Canberra, Australia, 2019).
[251] House of Representatives Standing Committee on Infrastructure and Communications: Inquiry into the use of subsection 313(3) of the Telecommunications Act 1997 by government agencies to disrupt the operation of illegal online services, "Balancing Freedom and Protection" (Canberra, Australia, 2015).

through social media, its effectiveness was monitored by its ability to observe its interventions in broader contexts. It judged its efforts most effective when they were taken up by mainstream media, and even, most effectively, in statements by US politicians. The seeding of such material was based on existing tendencies in US society. It worked by aggravating already existing grievances and amplifying them through emotional appeals, the spreading of fake news, and the skewing and distortion of issues, arguments and events.

Identification of malign Web activity

Identifying interventions by malign actors lies at the heart of identifying and countering mass influence campaigns. The Counter Narratives to Interrupt Online Radicalisation (CNIOR) project has identified netnographic analysis, a technique for the cultural analysis of social media and online community data, as a potential tool.[252] A netnographic study of the online activities of different groupings, and political and social orientations, may identify potential vulnerabilities and guard against their exploitation by malign actors.

Richards and Woods discuss the phenomenon of 'algorithmic drift', in which a site's algorithmic suggestions causes an Internet user to view more radical material and drift towards an environment containing more extreme messages. This can result in radicalisation and the creation of enclave communities reinforcing one other in their positions.[253] Such a process is clearly open to abuse by malign actors. Identifying such communities and the algorithms which reinforce them, in order to counter them, is a potential task of a counter-influence effort. If this approach were considered further research and consultation would be required to identify an appropriate ethical approach.

Fake news detection

Although the dissemination of fake news is not new, the phrase "fake news" was universally introduced following the 2016 US election. It has circulated widely – especially on social media – for many years. The most popular fake news articles were more widely shared on Facebook than the most popular mainstream news articles. Worse, many people who see fake news articles say they believe them.[254] Fake news misleads and manipulates how people interpret and react to real news.

Human fact checking cannot keep up with the volume of misinformation and speed at which it spreads. Also, we cannot have perfect machine learning algorithms. This is why Facebook prefers imperfect algorithms that allow false negatives to strict algorithms that return false positives.[255]

Techniques related to automatic fake news detection work as follows:

— features are extracted from news articles or posts;
— machine learning models are built to distinguish false or true content based on the features extracted.

We can classify these techniques based on how features are extracted:

1. **User-based:** User-based features represent the characteristics of those users who have interactions with the news on social media. These features can be categorised across different levels: Individual level and Group level.
2. **Post-based:** Post-based features focus on identifying useful information to infer the veracity of news from various aspects of relevant social media posts.

---

[252] Robert Kozinets, Pierre-Yann Dolbec and Amanda Earley, "Netnographic Analysis: Understanding Culture through Social Media Data," in *Sage Handbook of Qualitative Data Analysis*, ed. Uwe Flick (London: Sage, 2014), 262-272.
[253] Imogen Richards and Mark Woods, "Legal and Security Frameworks for Responding to Online Violent Extremism," in *The Handbook of Collective Violence: Current Developments and Understanding* (London and New York: Routledge, 2020). Book downloadable as a PDF via UNSW Canberra Academy Library.
[254] Craig Silverman and Jeremy Singer-Vine, "Most Americans Who See Fake News Believe It, New Survey Says," *Buzzfeed News*, 6 December 2016. https://www.buzzfeednews.com/article/craigsilverman/fake-news-survey.
[255] Natali Ruchansky, et al., "CSI: A Hybrid Deep Model for Fake News Detection," *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, (2017). https://dl.acm.org/doi/10.1145/3132847.3132877.

3. **Linguistic-based:** Linguistic-based features are extracted from the text content in terms of document organisations from different levels, such as characters, words, sentences and documents. Common linguistic features are often used to represent documents for various tasks in natural language processing.
4. **Network-based:** Network-based features are extracted via constructing specific networks among the users who published related social media posts.

Various systems and models utilising all four approaches already exist.

To what extent is fact checking an effective tool in countering mis- or disinformation and other malign activity?

In response to growing criticism over the rising tide of mis- and disinformation on its platform, in 2016 Facebook initiated a program to review material in conjunction with third-party fact checkers. Facebook content can be labelled as false, altered, partly false, missing context, satire or true. Identification of these can be met with three responses: removal, as in the case of hate speech, terrorist content or incitement to violence; reducing its distribution within the Facebook News Feed; or the application of warning labels and notifications. Facebook contends it does not allow ads that contain content "debunked by third-party fact checkers, including content rated as false, partly false, altered or missing context". Pages with content labelled as false or altered may see their distribution reduced, lose their ability to register as a News page and lose the ability to monetarise or advertise on the platform. Publishers may issue a correction or dispute a rating. Once misinformation has been identified, AI is used to find near identical versions on Facebook and Instagram, which are then similarly acted upon.[256]

Walter et al. report inconsistencies in the findings of research measuring the effectiveness of fact checking. Some studies[257] find that exposure to fact checking reduces the dissemination of misinformation, whereas others[258] find no such effect. Dias and Sippett note, however, by way of a caveat and criticism, that studies on the effects of fact checking have mostly been carried out in Western, English-speaking countries such as the US, UK and Australia.[259] There is, therefore, an absence of cross-cultural comparisons.

Mohseni et al. provide the following comments on fact checking and other misinformation counter measures:[260]

— Expert-review fact-checking methods are time-consuming, expensive and not scalable for stopping the spread of fake content in social media.[261]
— Machine learning methods are available to analyse (false) context. The linguistic features and writing styles of content can be analysed in order to detect possible false content. Spam detection and satire news detection is also available through machine learning algorithms.
— Clickbait detection algorithms are capable of analysing (in)consistencies between headlines and the content of news items as a form of fake news detection.
— Images and videos may also be evaluated or used as evidence. In the case of forged images and videos, researchers use deep learning methods to detect falsified content.

[256] "Fact Checking on Facebook," *Facebook for Business*: https://www.facebook.com/business/help/2593586717571940.
[257] Fridkin, Kim, Patrick J. Kenney, and Amanda Wintersieck, "Liar, liar, pants on fire: How fact-checking influences citizens' reactions to negative advertising." *Political Communication* 32, (2015): 127-151. doi:10.1080/10584609.2014.914613; Chung, Myojung, and Nuri Kim, "When I learn the news is false: How fact-checking information stems the spread of fake news via third-person perception." *Human Communication Research*. 47, no. 1 (2021): 1-24. https://doi.org/10.1093/hcr/hqaa010.
[258] Nyhan, Brendan, and Jason Reifler, "When corrections fail: The persistence of political misperceptions," *Political Behavior* 32 (2010): 303-330. doi:10.1007/s11109-010-9112-2.
[259] Nicholas Dias, and Amy Sippitt, "Researching Fact Checking," 605-613.
[260] Adapted from Mohseni, Sina, Eric D. Ragan, and Xia Hu, "Open Issues in Combating Fake News: Interpretability as an Opportunity," *Computer Science: Social and Information Networks*. (2019):1-7. p. 2. [1904.03016] Open Issues in Combating Fake News: Interpretability as an Opportunity (arxiv.org).
[261] See also Calum Thornhill, et al. "A Digital Nudge," 1-9. In addition, Lee, Dave. "Key fact-checkers top working with Facebook." *BBC News* 2 Feburary 2019. Available online at: https://www.bbc.com/news/technology-47098021 reports that some fact checking companies have withdrawn from working with Facebook.

— Natural language algorithms are being developed to distinguish between fake and genuine news content,[262] as well as fake news and satire.[263]

In January, 2018 Facebook announced it would consider the trustworthiness of news sources in its News Feed algorithm, based on a number of user surveys. As Sherman has pointed out, however, such an approach does not empirically gauge the accuracy of sources but rather the kind of news users want to read. Users who trust a source such as *Fox News* more than the *New York Times* are likely to see more posts citing *Fox News*. After analysing Facebook users' engagements with stories from a variety of sources, he concluded that sources such as the sensationalist *Daily Wire* receive more likes, shares and comments on Facebook than do sources such as *The Washington Post*, and that sources with more provocative headlines achieve higher visibility than do more reputable sources.[264]

In 2020 Facebook established the Facebook Oversight Board, dubbed by some Facebook's "Supreme Court", to allow an independent body of qualified individuals to act as a final arbitrator of content on Facebook.[265] While Facebook has spruiked the board consisting of members who have "lived in over 27 countries, speak at least 29 languages", critics have been sceptical about the ability of such a board to oversee the enormous task of fact checking the Facebook platform.

According to Walter et al., while fact checking can be effective in strengthening pre-existing convictions,[266] it is less effective correcting *misinformation* when it challenges one's existing beliefs.[267] Vedejová and Čavojová found that people tend to adopt a confirmatory strategy that aligned with their position.[268] They also found evidence of bias in the *interpretation* of information. This, however, was confined to more polarising topics (e.g., stem cell research and capital punishment). They concluded that people were more likely to twist evidence to fit their beliefs and theories in relation to issues important to their identity. When this was not the case, they were more open to persuasion by new evidence. They also noted that, in the case of extreme confirmatory bias,[269] when individuals see only what they want to see – in order to protect and maintain their entrenched beliefs, ideology and self-identity – debate becomes more difficult, and polarisation and extremism more likely.

---

[262] Choraś, Michał, Konstantinos Demestichas, Agata Giełczyk, Álvaro Herrero, Paweł Ksieniewicz, Konstantina Remoundou, Daniel Urda, Michał Woźniak, "Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study," Applied Soft Computing Journal, 101 (in press 2021); Gravanis, Georgios, Athena Vakali , Konstantinos Diamantaras, and Panagiotis Karadais, "Behind the cues: A benchmarking study for fake news detection," Expert Systems and Applications 128 (2019): 201-213; Oshikawa, Ray, Jing Qian, William Yang Wang. "A Survey on Natural Language Processing for Fake News Detection," Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020), Marseille, 11−16 May 2020: 6086−6093; Shariatmadar, David, "Could language be the key to detecting fake news?" The Guardian, 2 September 2019. Retrieved 19 March 2021 from: Could language be the key to detecting fake news? | Language | The Guardian; and Torabi Asr, Fatemeh, "The language gives it away: How an algorithm can help us detect fake news," The conversation 5 August 2019. The language gives it away: How an algorithm can help us detect fake news (theconversation.com) 19th March 2021.

[263] Levi, Or, Pedram Hosseini, Mona Diab, and David Broniatowski, "Identifying Nuances in Fake News vs. Satire: Using Semantic and Linguistic Cues," Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda, 2019: 31-35. 10.18653/v1/D19-5004. See also Zhang, Xichen, Ali A. Ghorbani, "An overview of online fake news: Characterization, detection, and discussion," Information Processing and Management 57, no. 2 (2020): 1-26 for a detailed fake news typology and detection overview).

[264] Len Sherman, "Zuckerberg's Promises Won't Fix Facebook, but You Can," *Forbes*, May 23, 2018: https://www.forbes.com/sites/lensherman/2018/05/23/zuckerbergs-promises-wont-fix-facebook-but-you-can/?sh=74149dee1e38.

[265] Nick Clegg, "Welcoming the Oversight Board," *Facebook*, May 6, 2020: https://about.fb.com/news/2020/05/welcoming-the-oversight-board/.

[266] See Kim, Antino, Patricia L. Moravec, and Alan R. Dennis, "Combating Fake News on Social Media with Source Ratings: The Effects of User and Expert Reputation Ratings," *Journal of Management Information Systems* 36, no. 3, (2019): 931-968. DOI: 10.1080/07421222.2019.1628921.

[267] See also Nieminen, Sakari, and Lauri Rapeli, "Fighting Misperceptions and Doubting Journalists' Objectivity: A Review of Fact-checking Literature," *Political Studies Review* 17, no. 3 (2019): 296-309; Young, Dannagal G. Kathleen H. Jamieson, Shannon Poulsen, and Abigail Goldring, "Fact-Checking Effectiveness as a Function of Format and Tone: Evaluating FactCheck.org and FlackCheck.org," *Journalism & Mass Communication Quarterly* 95, no. 1 (2018): 49-75.

[268] Vedejová, Dáša, and Vladimíra Čavojová, "Confirmation bias in information search, interpretation, and memory recall: evidence from reasoning about four controversial topics," *Thinking & Reasoning* (in press, 2021): 1-29. DOI: 10.1080/13546783.2021.1891967.

[269] A psychological process whereby people are more likely to search for and believe information that aligns with their pre-existing views.

Fact checking involves checking 'after-the-fact'. Thus, as Cazalens et al. point out, refutation of false content may come too late to prevent rumours and false information from taking hold in the minds of those for whom such claims sit well. In fact, research indicates that mis-/disinformation may be difficult to counter when it resonates effectively with people's emotional commitment to a position. *Belief perseverance* refers to the phenomenon of people refusing to change their position in the light of new information, also known as the *backfire effect*.[270]

What makes an effective fact checking workforce?

Facebook's fact checking initiative has been criticised for allegedly providing insufficient resources to its third-party fact checkers to cover a platform with more than 2 billion users.[271] A study by the Massachusetts Institute of Technology (MIT) Sloan School of Management also criticised the rating system for the "implied truth effect" that creates the impression that, as some posts are rated as false and misleading, those that are not should be regarded as reliable. The reality is that fact checkers are capable of addressing only a tiny fraction of the disinformation on Facebook.[272]

A vulnerability in Facebook's efforts to counter mis- and disinformation is its reliance on outsourced, third-party fact checking services staffed by often underpaid human operators around the world who are often under-resourced and overwhelmed by the sheer volume of material. Nevertheless, the combination of machine learning algorithms *and* human fact checkers has been found to increase the effectiveness of fact checking.[273] There is also some indication that fact checking, and the provision of warning labels and alternative material on Facebook may have some impact.[274] A study found that while false content continued to rise on Twitter between 2016 to July, 2018 it fell sharply on Facebook during this period. Researchers concluded that Facebook's efforts to limit misinformation after the 2016 US presidential elections may have had a meaningful impact.[275]

Ghosh and Scott contend that strategies to counter disinformation must consider a number of digital marketing features, including:

— Behavioural data tracking that enables precision targeting;
— Online ad buying to reach and impact certain audiences;
— Search engine optimisation that tricks algorithms and dominates search results;
— Social media management services that preconfigure messages for select audiences across multiple media channels; and
— Marketing AI that improves behavioural data tracking, audience segmentation, message targeting and campaign management.[276]

---

[270] See Brendan Nyhan and Jason Reifler, "Displacing Misinformation about Events: An Experimental Test of Causal Corrections," Journal of Experimental Political Science, 01 April 2015, https://doi.org/10.1017/XPS.2014.22; Brendan Nyhan and Jason Reifler, "When corrections fail: The persistence of political misperceptions," *Political Behavior* 32, no. 2 (2010): 303-330; Brendan Nyhan, and Jason Reifler, "Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information," *Vaccine* 33, no. 3 (2015): 459-464. See also Jiang, Shan, and Christo Wilson "Linguistic Signals under Misinformation and Fact-Checking: Evidence from User Comments on Social Media," Proceedings of the ACM on Human-Computer Interaction 2, No. CSCW, Article 82 (November 2018): 1-23; and Kathryn Haglin. "The limitations of the backfire effect," Research and Politics, July-September (2017): 1-5. The limitations of the backfire effect (sagepub.com) for a challenge to the "backfire effect." Also, Mazarr, Michael J., Ryan Bauer, Abigail Casey, Sarah Heintz, and Luke J. Matthews, "The Emerging Risk of Virtual Societal Warfare: Social Manipulation in a Changing Information Environment", Santa Monica, CA: RAND Corporation, 2019. https://www.rand.org/pubs/research_reports/RR2714.html claim that the backfire effect is contested and may be specific to certain situations rather than generalisable.
[271] Chris Mills Rodrigo, "Critics fear Facebook fact checkers losing misinformation fight," *The Hill*, 20 January 2020: https://thehill.com/policy/technology/478896-critics-fear-facebook-fact-checkers-losing-misinformation-fight.
[272] Peter Dizikes, "The catch to putting warning labels on fake news," *MIT News*, March 2, 2020: https://news.mit.edu/2020/warning-labels-fake-news-trustworthy-0303.
[273] Conversation with A/Prof Stephen Doherty.
[274] The services of many fact checkers are acquired through Amazon's crowd sourcing website, *Mechanical Turk*.
[275] Hunt Allcott, Matthew Gentzkow, and Chuan Yu, "Trends in the Diffusion of Misinformation on Social Media," Stanford Institute for Economic Policy Research, October 2018. fake-news-trends.pdf (stanford.edu).
[276] digital-deceit-final-v3.pdf (d1y8sb8igg2f8e.cloudfront.net).

Other approaches

In 2008, Weimann and von Knop argued that malign actors create noise by the deliberate propagation of misinformation, in order to undermine public confidence in the government's policies, and aggravate pre-existing fears and neuroses in sections of the population.[277] Research by the Counter Narratives to Interrupt Online Radicalisation (CNOIR)[278] project indicates that such noise may be effective in countering online radicalisation. To be most effective, such noise should be undertaken by people who are "highly connected in their respective social networks", and include a discussion of "broader community concerns and priorities", rather than simply focussing on the extremist material the noise is engaging with.[279]

Technology

As previously discussed, an Australian counter-influence capability will need to be agile and ready to respond to technological changes and developments, whether implemented by Facebook and other tech companies, or by potentially malign actors who may use technological developments to their advantage. That said, discussions with supporting academics indicate that the technology required for the provision of such a capability should, at present, include a number of capabilities, such as:

— detecting automated behaviours;
— Indexing images – detecting the reusing, repurposing and manipulation of images for potentially malign purposes;
— the ability to detect unusual changes, such as when an account has been taken over;
— the correlation of behaviour between accounts to try and identify botnets and manipulated accounts; and
— the examination of behaviour across platforms to detect potentially malign activity.

Advice indicates that, due to Facebook privacy constraints, at present, it is easier to collect this information on Twitter. It has been suggested that such factors could be first detected on that platform and the information used to determine whether these activities are also occurring on Facebook. The next tranche could further explore the specific technologies required to identify malign Web activity, as well as the partnerships that may also be required to most effectively do this.

Other matters to be further explored include:

— The utilisation of message data tracking. If a message is posted on a platform, it is possible to trace it and determine how it is shared. If it is shared to a reasonable extent, it will create a "tree" of the users, which can then be used to determine the characteristics of the tree, those participating in it and the online community the message engages in. This in turn can be used to detect the activities of online malign actors.
— The proposed capabilities of an Australian counter-influence capability in regard to messaging, the sending out of messages to groups, the role the centre should play in this and the impact such factors are likely to have.
— The potential use of cookies in terms of what information can be garnered from them and how it could be used.
— The potential use of Web scraping to obtain information.
— The use of AI, as used by Facebook, to locate disinformation on other pages once it has been identified and countered.

[277] Gabriel Weimann and Katharina von Knop, "Applying the Notion of Noise to Countering Online Terrorism," *Studies in Conflict & Terrorism*, 31, no 10 (2008): 883.

[278] Carmen Jacques and Kosta Lucas, *Counter Narratives to Interrupt Online Radicalisation (CNIOR) project proposal.* CNIRO internal document. The UNSW team is grateful to Carman Jacques and Kosta Lucas of the CNIOR project for the provision of information and sources.

[279] Talene Bilazarian, "Countering Violent Extremist Narratives Online: Lessons From Offline Countering Violent Extremism," *Policy & Internet*, 12, no. 1 (May 2019): 46. https://onlinelibrary.wiley.com/doi/abs/10.1002/poi3.204.

— The potential use of sets of banded algorithms to use with different groups and in different settings to determine which is most effective. Banded algorithms are used by Facebook and it is possible that an operations centre will be able to use them too, if in a situation that is much more constrained regarding data on Facebook.

— The establishment of Facebook accounts for the purposes of locating and countering disinformation and malign activities.

These areas are problematic not only in terms of the technology required to enact them but also in terms of what ethical parameters should be established regarding their use. The latter will require a broader discussion with relevant parties over the longer term, longer than available in this preliminary study.

Nemr and Gangware identify three broad knowledge gaps and technology challenges that impact on efforts against disinformation:

1. Technology gaps. Arguably an arms race in which researchers, technologists and governments scramble to develop tools to detect, counter, and keep pace with malicious actors' methods and activities.

2. Structural challenges. Economic incentives for developing counter information technology, the dearth of available data sets to train machine tools, the slow rate of adaption of existing tools.

3. The gap in understanding how technologies such as AI are evolving and, because of this, the threat of disinformation.[280]

More specifically, they identify the following challenges:

— **Bots.** Bots are becoming more sophisticated to exploit advertising markets. They make the point that "Detecting spam box on Twitter based on syntax, semantics, or network creatures is effective", but that the next generation of bots may be more difficult to detect or counter as they become more intelligent.

— **Photos and videos**. Detecting altered photos and videos at scale is difficult, and rapidly advancing AI and deep learning technology is making synthetic media easier to produce. The US Defence Advanced Research Projects Agency's "media forensics" program is working to develop and deploy tools to automate the assessment of an image's integrity.

— **Fact checking.** Automated fact checking is a new area of development. Ongoing hurdles include how to teach computers to identify the parts of a sentence that should be fact checked. Nevertheless, it is an area likely to increase in importance in the coming years.

— **Structural challenges.** The sheer magnitude of content and platforms is one of the biggest obstacles.

— **Encryption.** A problem. Disinformation that spreads through private encrypted chat groups cannot be easily detected through technical means, nor fact checked or countered.

— **Verification tools.** While a limited number of automated social verification tools exist, their use remains relatively limited.

— **Access to data.** Researchers lack the access to data necessary to train machine learning tools to operate in countering these information efforts. Companies, particularly Facebook, rarely share such data. Machine learning is therefore insufficient and needs to be accompanied by human review.

---

[280] Christina Nemr and William Gangware, *Weapons of Mass Distraction: Foreign State-Sponsored Disinformation in the Digital Age* (Plymouth, Minnesota: Park Advisors, 2019), 39-43.

Theme 4: Campaign Awareness and Sensemaking

Research Question 4: How is Facebook able to achieve and maintain awareness of the impact of its influence activities?

Since its development beyond a student-only platform in 2006, Facebook has empowered people to share and even publicise personal information, preferences and other views like never before and, in doing so, revolutionised the manner and extent to which people are able to connect worldwide. Such a transformation in 'community' engagement is in keeping with Facebook's *original* mission statement which reads: *To give people the power to share and make the world more open and connected*.

As noted previously, getting people to engage with each other online as prosumers is a key component of Facebook's business model and a way of achieving the business goal of its influence activities. Given this, addressing the questions below will help in addressing Research Question 4.

Given Facebook's goal, how does it monitor its success in achieving this goal?

Facebook's attempt at transparency – that is, making available information about what it is capable of monitoring, in terms of user engagement and its effects (say, in relation to responses to advertisements) – reveals *something* about the organisation's ability to monitor and record user engagement, as well as what it considers to be of *interest* and/or *value* to its users and third-party customers. After all, where there is an absence of legal requirement (although perhaps not public pressure, see below), why does Facebook choose to make available (to be transparent about) the information it does? To illustrate: In addition to recording the time someone spends on its platform, Facebook provides engagement metrics on (*inter alia*) the number of 'likes' given, photos uploaded, events currently attending, groups joined, photos tagged, links and questions posted, and status and location (i.e., check-in) updates.[281] The purpose of this transparency, one might think, is to make salient an individual's social presence on the platform and their (Facebook) popularity (as discussed earlier).

Information available to those who advertise on Facebook (about their advertisement posts) is available via *Facebook insights* or *analytic tools* and includes tracking metrics of the kind found in Table 3.

Table 3: Examples of information Facebook monitors [282]

Tracking metrics	Description (based on advertisement)
Engagement	The number of actions – 'likes', shares, comments – taken.
Reach	The number of people who have seen the advertisement.
Impression	The number of times the advertisement has been seen and/or acted on by the *same* person.
Referral traffic	The number of visits to the advertiser's own website, via Facebook.

---

[281] Tasos Spiliotopoulos, and Ian Oakley, "Understanding Motivations for Facebook Use: Usage Metrics, Network Structure, and Privacy," *CHI 2013*, Paris, France (April 27-May 2, 2013): 1-10.
[282] Dominique Jackson, "11 Facebook Metrics Every Brand Needs to Track," *Spout Blog*, 7 August 2020. https://sproutsocial.com/insights/facebook-metrics/.

This information is of particular use to Facebook advertisers given Yoon et al.'s finding that increased engagement with online advertisements – commenting positively, rather than just clicking 'like' or even sharing – correlates with higher company revenue for advertisers.[283] Yoon et al. conjecture that the one click – or shallow – engagement required to 'like' or share a post, does not reflect the same commitment and, therefore, engagement with the product a (positive) comment does. The information provided by Facebook is no doubt of benefit, then, to advertisers when they are assessing the effectiveness of their campaign. But it also benefits Facebook, in terms of demonstrating, and thereby promoting, the value of its big data and predictive algorithms at compiling a target audience for advertisers (i.e., showing when the approach taken is working and when it is not). It also informs Facebook about the effectiveness of its techniques of influence and its use of persuasive technology (i.e., the effectiveness of presenting an advertisement in the same, or very similar, way a post is shared by a 'friend').

According to Flyverbom, however, Facebook's transparency indicators (such as those discussed so far; also see below) are best understood in conjunction with how Facebook (as well as other forms of digital technology) has fundamentally altered the way we present, interact with and record information. In this context, transparency does not exist simply to provide insight and clarity but to *mediate* and *manage visibilities*.[284]

A recent example of this can be seen in the public scrutiny of the ongoing role played by Facebook in disseminating political mis- and disinformation, which was felt to be particularly acute during the 2016 US presidential election (see Case Study 2 on *Cambridge Analytica*). In response to this scrutiny, Facebook's *Page Transparency* was created to make available disclosure information such as:

— The date a particular page was created.
— The primary country where the page is managed and the number of people who manage it.
— Any previous name changes or page mergers.
— The confirmed business or organisation that claims ownership of the page and any additional businesses or organisations that have been granted access to help manage it.
— Whether the page belongs to a state-controlled media organisation.[285]

The means and extent of Facebook's transparency is also assessed by the Transparency Advisory Group which has offered recommendations such as:[286]

— Prioritising the release of accuracy rates for both human and automated decisions regarding post removals that violate Facebook's Community Standards (see Section 4.4).
— Releasing the reversal rate after appeal of the decision to remove posts.
— Share the inter-rater reliability of human reviewers concerning decisions to remove posts.[287]
— Check reviewers' judgments not only against an internal 'correct' interpretation of the Community Standards, but also against a survey of users' interpretations of the Standards.

It is also important to note that there remains much that Facebook is not transparent about. In 2018, for example, after the Cambridge Analytica scandal, Facebook introduced two additional transparency features. *Download Your Information* is said to provide users with information about businesses that have acquired user details from Facebook. *Off-Facebook Activity* is supposed to allow users to see and control data about themselves provided to Facebook by other websites and apps. *Privacy International*, however, reports that, in each case, the information is

---

[283] Gunwoo Yoon, Cong Li, Yi (Grace) Ji, Michael North, Cheng Hong, and Jiangmeng Liu, "Attracting comments: Digital engagement metrics on Facebook and financial performance," *Journal of Advertising* 47, no. 1 (2018): 24-37.
[284] Mikkel Flyverbom, "Transparency: Mediation and the management of visibilities," *International Journal of Communication* 10 (2016): 110-122.
[285] https://www.facebook.com/help/323314944866264 See also Matthew Binford et al., "Invisible transparency," 70-83 for a useful discussion on the introduction of disclaimers on US political advertising, as well as evidence suggesting that these disclaimers are very often not noticed by users.
[286] Adapted from Ben Bradford, Florian Grisel, Tracey L. Meares, Emily Owens, Baron L. Pineda, Jacob N. Shapiro, Tom R. Tyler, and Danieli E. Peterman, "Report of the Facebook Data Transparency Advisory Group," *The Justice Collaboratory: Yale Law School*: 1-43.
[287] Inter-rater reliability is the measure of agreement between raters (in this case reviewers), typically measured as a correlation coefficient.

limited to the company's name only. No contact details or indication of what these companies know is provided. *Privacy International* has described Facebook's response to the call for greater transparency and privacy protection, even when taking into account international differences, as "a tiny sticking plaster on a much wider problem".[288]

Are there unintended consequences of Facebook's influence activities?

On Facebook, people regularly share – i.e., make available to others – their thoughts, feelings and behaviours. It is therefore true, insofar as people have *access* to this information more readily and in greater quantities than ever before, that the world is a more open and connected place, just as Facebook aspired to make it in their original mission statement. It seems less reasonable to claim, however, that the world, or even the world of Facebook, has become a more communal and unified place as a result. Zuckerberg appears to have come to the same conclusion when he states:

> Social media is a short-form medium where resonant messages get amplified many times. This rewards simplicity and discourages nuance. At its best, this focuses messages and exposes people to different ideas. At its worst, it oversimplifies important topics and pushes us towards extremes…
>
> …If [polarization] continues and we lose common understanding, then even if we eliminated all misinformation, people would just emphasize different sets of facts to fit their polarized opinions.[289]

What Zuckerberg and others have come to realise, it seems, is that Facebook's original mission statement is flawed because although the aim of giving people more power to share, and create a more open and connected world, is achievable, it does not follow from this that the outcomes will automatically be positive. It is a point Anti-Defamation League CEO Johnathan Greenblatt has made repeatedly in his ongoing criticism of Facebook's (alleged) poor record of tackling hate speech on its platform. He states: "We've been at this work of fighting anti-Semitism and bigotry in all forms for over 100 years… And frankly, we believe that Facebook is the front line in fighting hate today."[290]

Kelly, likewise, points out that the act of connecting people based on shared interests does not always guarantee a good outcome.[291] After all, racists and homophobes can thrive in the kind of environment of like-minded individuals found on Facebook.[292] Moreover, Horwitz and Seetharaman claim that, in 2016, Facebook was made aware that not only were its algorithms creating more divisiveness on the platform, and even exploited this to increase user engagement, they were also responsible for the growth of extremist groups on Facebook. A total 64% of all extremist group 'joins', Facebook was told, resulted from its recommendation tools (e.g., "Groups You Should Join").[293]

[288] https://privacyinternational.org/long-read/3372/no-facebook-not-telling-you-everything.
[289] Mark Zuckerberg, "Building Global Community," Facebook, May 6, 2021. https://de-de.facebook.com/notes/mark-zuckerberg/building-global-community/10103508221158471.
[290] Cited in Shirin Ghaffray, and Rebecca Heilweil, "Why Facebook is "the front line in fighting hate today," *Recode*, 15 July 2020, p.1. Retrieved 2 March 2021 from: https://www.vox.com/recode/2020/7/15/21325728/facebook-stop-hate-for-profit-campaign-jonathan-greenblatt-anti-defamation-league.
[291] Heather Kelly, "Mark Zuckerberg explains why he just changed Facebook's mission," *CNN Business*, 22 June 2017. Retrieved 1 March 2021 from: https://money.cnn.com/2017/06/22/technology/facebook-zuckerberg-interview/index.html?iid=EL.
[292] Jemima McEvoy, "Study: Facebook Allows And Recommends White Supremacist, Anti-Semitic And QAnon Groups With Thousands Of Members," *Forbes* (4 August 2020). Retrieved 1 March 2021 from: https://www.forbes.com/sites/jemimamcevoy/2020/08/04/study-facebook-allows-and-recommends-white-supremacist-anti-semitic-and-qanon-groups-with-thousands-of-members/?sh=1d0b993e6bbd. See also https://www.businessinsider.com.au/facebook-civic-groups-internal-docs-capitol-riot-misinformation-hate-2021-2?r=US&IR=T.
[293] Jeff Horwitz, and Deepa Seetharaman, "Facebook Executives Shut Down Efforts to Make the Site Less Divisive," Wall Street Journal, 26 May 2020. Retrieved 18 March 2021 from Facebook Executives Shut Down Efforts to Make the Site Less Divisive - WSJ.

At the same time as Facebook is using artificial intelligence, including 'hate speech algorithms'[294] and teams of human monitors, to identify individuals and groups that are using the platform to promote a terrorist agenda, including White Supremacy, Zuckerberg continues to insist Facebook is not in the business of policing content and is, instead, a guardian of free speech, even when the material is objectionable.[295] Lately, however, Zuckerberg's position on this seems more nuanced, something we will now discuss.

Facebook: A platform or a publisher?

From 2008 to 2018, Facebook's default response to demands that it should police content posted on its platform or face legal consequences was to claim that it was a platform and not a publisher.[296] It should therefore not interfere with content, not only to remain neutral (maintain "platform neutrality") but also to avoid being misconstrued as a publisher, which would result in the company becoming liable for content published on its platform (according to US legislation).[297] During the 2018 US Senate hearing, however, Zuckerberg indicated that editorial decisions were being made about content in accordance with Facebook's *Community Standards* (again, see Section 4.4).[298] Around the same time, although in a different context, Facebook's legal representatives acknowledged that Facebook was indeed a publisher.[299] Zuckerberg nevertheless expressed resistance in 2019 to the idea that Facebook should act as the "truth police" when it came to content posted on its platform and, subsequently, announced that Facebook would take a hands-off approach to policing material not clearly violating its standards.[300] Disagreeable content should therefore be allowed as long as it is not hate speech or does not, in some other way, cross the line.[301] In 2020, however, and perhaps building on this last point, he conceded that more moderation of content on the Internet was required, although this should be based on a universally applicable framework.

Since 2018, Facebook appears to have shifted its position regarding the management of content on its platform, or has at least sought to clarify it publicly. Yet the nature of its clarification suggests that Facebook both is and is not a publisher, or at least is 'kind of' a publisher, insofar as it makes editorial decisions about content. Facebook's position is, however, entirely consistent with the editorial freedoms afforded it as a *platform* (not a publisher) by the amended Section 230(c)(1) of the 1996 US *Communications Decency Act* (CDA). According to this Section, "no provider or user of interactive computer services shall be treated as the publisher or speaker of information provided by another information content provider".[302]

Section 230 removes any liability from *platforms* that publish false, dangerous, or misleading content online,[303] and is perhaps intended to reduce the likelihood that social media platforms, through fear of litigation, will become overly restrictive when it comes to the content they will allow to be posted. Such overly restrictive behaviour would curb the empowering effect one-to-many communications and minority participation has had on free speech, which social media platforms have enabled.[304] It would also be extremely difficult (if not impossible) for social media platforms to police all posted content in real time. Nevertheless, to encourage social media platforms to take some

---

[294] https://www.adl.org/news/press-releases/adl-statement-on-facebook-reportedly-overhauling-hate-speech-algorithms.

[295] See also Andre Oboler, "The rise and fall of a Facebook hate group," *First Monday 13,* no. 11 (2008): https://doi.org/10.5210/fm.v13i11.2254. See also https://www.adl.org/news/press-releases/after-long-fight-adl-is-relieved-at-facebook-announcement-that-it-will-remove for interesting recent development in relation to Anti-Semitism.

[296] Can Facebook ever really be neutral? — Quartz (qz.com); https://www.reuters.com/article/us-facebook-zuckerberg-idUSKCN1141WN; Facebook's and YouTube's 'Platform' Excuse Is Dying - The Atlantic.

[297] Platform neutrality is synonymous with platform passivity: the idea that platforms are or should be simply a means by which content is expressed, and therefore a vehicle for this expression, much in the same way that a telephone company provides a means of communication and is not responsible for what is said on the telephone. For a detailed discussion and criticism of the idea of platform neutrality, see Anupam Chander, and Vivek Krishnamurthy, "The myth of platform neutrality," *2 Georgetown Law Technology Review 400*, (2018): 400-416.

[298] https://www.facebook.com/communitystandards/.

[299] Is Facebook a publisher? In public it says no, but in court it says yes | Facebook | The Guardian.

[300] Horwitz and Seetharaman, "Facebook Executives Shut Down Efforts."

[301] Heather Kelly, "Mark Zuckerberg explains why he just changed Facebook's mission."

[302] Taken from Caitlin McKeown, "Facebook, defamation, and terrorism: Who is responsible for dangerous posts on social media?" *Tulane Journal of International and Comparative Law* 26 (2017): 163-187, p. 167.

[303] Taken from McKeown, "Facebook, defamation, and terrorism," 163-187.

[304] Chander and Krishnamurthy, "The myth of platform neutrality," 400-416.

responsibility for the content posted on their platforms, without fear of prosecution, Section 230 forbids the imposition of *publisher liability* on a service provider for the exercise of its *editorial and self-regulatory functions*.[305] In other words, social media platforms like Facebook can exercise editorial control without being construed as a publisher and, therefore, without being liable for any and all content published on its platform simply because it has made editorial decisions about *some* content.

To achieve these two aims – i.e., continuing to facilitate freedom of speech whilst encouraging the editorial management of content – Section 230 recognises a distinction between types of Internet service providers. An *interactive computer service*, which includes Facebook, is "any information service system, or access software provider that provides or enables computer access by multiple users to a computer server".[306] An *information content provider*, on the other hand, is "any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service".[307] This means that an individual or group generating content on their Facebook page would count as an information content provider and, in accordance with CDA, be liable for the content posted, whereas Facebook, as an interactive computer service, would be immune from prosecution for the content posted by this group (the information content provider).

For Goldman, Section 230 "is a flagship example of *Internet exceptionalism*" because it allows the Internet to be regulated differently to other media.[308] By way of a small dent in this alleged exceptionalism, however, in 2018, then President Trump signed into law the US Congress *Fight Online Sex Trafficking Act* (FOSTA) and the Senate *Stop Enabling Sex Traffickers Act* (SESTA). Each act removes the immunity granted to platforms under Section 230 of the CDA when dealing with sex trafficking, making Facebook potentially responsible for such material posted by users.[309] More recently (early 2021), senior House Republican, Jim Banks, introduced a bill seeking to ensure that social media platforms are not protected from prosecution – through the enforcement of Section 230 – when they knowingly share illicit content, such as child pornography, posted by users on their platform.[310]

Perhaps in response to these (and proposed) changes to legislation, Zuckerberg seems willing to concede the need for reform and openly favour striking a balance between not overzealously acting as the thought police, and not failing in one's social responsibility to manage content. In a recent 2020 US Senate hearing, for example, he went as far as to suggest amending Section 230 of the CDA so that social media platforms are required to have adequate systems in place to address unlawful content.[311] This suggestion was possibly motivated by reports that President Joe Biden's administration is in favour of reforming Section 230.[312]

The seeming change in Facebook's approach to policing content can be traced to at least 2017 when Facebook's mission statement changed. The revised statement now reads: "To give people the power to build community and bring the world closer together."[313] It may be, of course, that the sentiment expressed within this new mission is what Zuckerberg had aspired to all along, even in the original statement. If so, then although this aspiration is more explicitly stated in the new version, the idea of a community based on shared understanding and an acceptance, or at least tolerance, of difference, as well as a desire to find common ground and help bring us *all* closer together is still, at best, implied. It is, however, a view more explicitly articulated by Zuckerberg elsewhere: "It's important to

---

[305] See Zeran v. American Online Inc. (harvard.edu).
[306] McKeown, "Facebook, defamation, and terrorism," 163-187, p. 168.
[307] McKeown, "Facebook, defamation, and terrorism," 163-187, p. 168.
[308] Eric Goldman, "Want to Kill Facebook and Google? Preserving Section 230 Is Your Best Hope." Balkinization: *New Controversies in Intermediary Liability Law,* (2019): 1-5.
[309] Aja Romano, "A new law intended to curb sex trafficking threatens the future of the internet as we know it," *The Verge*, 2 July 2018: https://www.vox.com/culture/2018/4/13/17172762/fosta-sesta-backpage-230-internet-freedom.
[310] Nihal Krishan, "Top House conservative introduces bill to overhaul Section 230 to curb child pornography," *Washington Examiner*, 18 March 2021.
[311] Jowitt, Tom, "Zuckerberg Proposes 'Thoughtful Reform' Of Section 230 Liability," *Silicon*, 25 March 2021.
[312] Makena Kelly, "Biden's Commerce nominee backs changes to Section 230," *The Verge*, 26 January 2021.
[313] Nick Statt, "Mark Zuckerberg just unveiled Facebook's new mission statement," *The Verge*, 22 June 2017.

give people a voice, to get a diversity of opinions out there, but on top of that, you also need to do this work of building common ground so that way we can all move forward together."[314]

Does the change in Facebook's mission statement reflect a realisation, perhaps in hindsight, that facilitating openness and connectedness, *alone*, can lead to the kinds of negative consequences that critics have accused Facebook of not doing enough to combat? One thing it seems reasonable to conclude is that Facebook's content moderation is an expression of the balance it tries to strike between its mission statement and the company's bottom line.[315] As Nick Clegg, Facebook Vice-Chairman of Global Affairs and Communication, noted in 2021: "It is not in Facebook's interest – financially and reputationally – to continually turn-up the temperature and push users towards ever more extreme content."[316] Given Facebook's recent public acknowledgement of the negative consequences of enhanced openness and connectedness, whether intended or simply foreseen, or neither, what mechanisms does Facebook *now* have in place to help it become aware of these and other consequences (whether good or bad) of its influence operations?

How is Facebook able to achieve and maintain awareness of any unintended consequences of its influence activities?

In response to recent pressure,[317] Facebook published a public-facing *Community Standards* document detailing the rules that govern Facebook usage (such as not posting hate speech or sexual content and nudity). In conjunction with this more recent act of transparency, Facebook regularly publishes its *Community Standards Enforcement Report*.[318] Contained within its pages is information on the number of identified cases of hate speech, bullying and harassment, and updates on how Facebook is helping to manage election integrity and combat misinformation. Facebook has also recently published the criteria that must be satisfied in order for a post to be removed from its platform.[319] Silow-Carroll, however, notes how decisions about removing posts can sometimes appear capricious, and be reliant on knowledge (or sometimes succumb to ignorance) of the local standards set within the numerous countries in which Facebook operates.[320] Wagner et al., for example, report how Facebook's *Community Standards* document does not always align with other countries' laws.[321] For example, Facebook's *Community Standards* is more tolerant than the *German Network Enforcement Act* when regulating hate speech. This has resulted in Facebook having to conform to the local regulations in Germany, even when content may not violate its own standards.

The extent of Facebook's global outreach means that it must navigate its way through different 'local' regulations. This makes any attempt at a centralised approach to policing/managing content – see discussion on Facebook's Oversight Board (sometimes referred to as the 'Supreme Court', Section 3.10) – complex and potentially fraught with difficulty.

Facebook regularly evaluates its metrics in order to estimate the number of "duplicate" and "false" accounts on its platform. The Facebook SEC submission estimates that in the fourth quarter of 2020 duplicate accounts may have represented 11% of MAUs and false accounts, 5%. Facebook defines "violating" accounts as those that are intended to be used for purposes that violate Facebook's terms of service, including bots and spam. In the fourth quarter of 2020 it estimated 3% of worldwide MAP accounts were in this category, and while Facebook does state that, "from time to time", it disables duplicate, false and violating accounts, the document (SEC submission) does not provide

[314] Nick Statt, "Mark Zuckerberg just unveiled Facebook's new mission statement."
[315] Presuel, Rodrigo, and Sierra, "Algorithms and the News," 261-285.
[316] https://nickclegg.medium.com/you-and-the-algorithm-it-takes-two-to-tango-7722b19aa1c2.
[317] Robert Gorwa, and Timothy G. Ash, "Democratic transparency in the platform society," In Social media and democracy: The state of the field, prospects for reform, edited by Nate Persily, and Josh A. Tucker, 286-312, (Cambridge: Cambridge university Press), 2020.
[318] https://about.fb.com/news/2021/02/community-standards-enforcement-report-q4-2020/ (Accessed 26 February 2021).
[319] https://www.facebook.com/communitystandards/ (Accessed 26 February 2021).
[320] Andrew Silow-Carroll, "How Facebook monitors harmful content: 5 takeaways," *The Times of Israel*, 29 December 2018. Retrieved 26 February 2021 from: https://www.timesofisrael.com/how-facebook-monitors-harmful-content-5-takeaways/.
[321] Ben Wagner, Krisztina Rozgonyi, Marie-Therese Sekwenz, Jennifer Cobbe, and Jatinder Singh, "Regulating transparency? Facebook, Twitter and the German Network Enforcement Act," *FAT* '20*, Barcelona, Spain (27-30 January 2020): 1-11.

figures in this regard.[322] In 2018, however, it announced that in the first quarter of the year it had shut down 583 million fake accounts.[323]

Given the importance of user authenticity to its targeted advertising model, Facebook has a clear incentive to police the authenticity of identity more than content. From a business perspective, it has a much stronger incentive to remove false identities than mis- or disinformation. The issue of 'fake news' on Facebook, however, is enormous, with a Pew Research Centre study in 2017 determining that just under half (45%) of US adults use Facebook for news, and half of these get news from Facebook alone. In contrast, news users of LinkedIn, Instagram, Snapchat and WhatsApp are more likely to obtain news from multiple sources.[324] In light of recent announcements by Zuckerberg regarding increased editorial intervention, Facebook seems to be moving towards achieving a greater balance between business benefits and social responsibility.

---

[322] Form 10-K, Facebook, Inc., 4-5.

[323] Len Sherman, "Zuckerberg's Broken Promises Show Facebook Is Not Your Friend," *Forbes*, May 23, 2018: https://www.forbes.com/sites/lensherman/2018/05/23/zuckerbergs-broken-promises-show-facebook-is-not-your-friend/?sh=14f114cb7b0a.

[324] Elizabeth Greco, "More Americans are turning to multiple social media sites for news," *Pew Research Center Factank*, 2 November 2017: https://www.pewresearch.org/fact-tank/2017/11/02/more-americans-are-turning-to-multiple-social-media-sites-for-news/.

Observations and Recommendations

1. An Australian counter-influence capability will likely require a strategy and framework for clear and transparent communication with the public. This includes the following considerations:
   (i) Guidelines and definitions for what constitutes ethical persuasion.
      - An under-developed research area that involves determining whether the influence Facebook (or other social media platforms) exerts on users amounts to a form of manipulation or non-rational persuasion (or possibly both, at different times and in different contexts).

   (ii) Capacity to differentiate between authentic and inauthentic patterns of social engagement in the context of true or false content.
      - Authentic or inauthentic accounts can create and share content, irrespective of its truth-value. Fake content, for example, can be created and shared for benign reasons and true content for nefarious reasons.

2. How to operate in a social media landscape of powerful distributions of production networks with concentrated ownership.
   (i) Advancement in machine learning and AI.
   (ii) Consider investment in the development of information analytics. Facebook is deliberately opaque in this regard. However, Australian centres of excellence could be leveraged to develop fit-for-purpose, sovereign analytic tools and techniques.
   (iii) Netnographic analysis – a technique for the cultural analysis of social media and online community data – is a potential tool to identify malign online actors and their behaviours and vulnerabilities.

3. A growing and diverse workforce.
   (i) Skills required to identify nefarious activity disguised as benign social engagement. A mix of technical and non-technical roles.
      - Whilst technology is advancing, Facebook's adapted processes for fact checking and content moderation reveal the limitations of algorithms to 'make the right choices'. Natural language algorithms are more effective in conjunction with human fact checkers when seeking to identify nefarious activity disguised as benign social engagement.

   (ii) Mix of military and non-military employees.
      - Facebook's future workforce projections of minimal human operators are most likely overestimated. An Australian counter-influence capability will similarly not be able to rely solely on automated algorithms. An Australian counter-influence capability will require a workforce trained in understanding the complexities of operating online, and have a mix of technical and non-technical roles.

   (iii) Workplace health and psychological support.
      - Given the evidence that habitual exposure to malign activity may cause trauma in participants, workers should be provided support services, including resiliency training and counselling. Examples of such support systems exist in organisations (other than Facebook) undertaking content moderation.

4. A Facebook liaison facility.

   (i)   Examples of government cooperation with Facebook exist and could be used as models for the establishment of a Facebook Liaison facility as part of an Australian counter-influence capability.

   (ii) Existing Facebook obligations under the Telco Act may provide the basis for the establishment of such a facility.

References

Aguirre, Diego. "Ten Things You Need to Know About Spark AR," *Diego Aguirre*, May 13, 2019.

Aloudat, Anas, Ons Al-Shamaileh, and Katina Michael. "Why some people do not use Facebook?" *Social Network Analysis and Mining* 9, no. 1 (article 19) (2019): 1-14. https://doi.org/10.1007/s13278-019-0564-z

Anagnostopoulos, Aris, Ravi Kumar, Mohammad Mahdian. "Influence and Correlation in Social Networks." *KDD '08: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, August 24–27, 2008, Las Vegas, Nevada, USA: 7-15.

https://doi.org/10.1145/1401890.1401897

Andersen, Jack., and Sille O. Søe. "Communicative Actions We Live By: The Problem with Fact-Checking, Tagging or Flagging Fake News - the case of Facebook." *European Journal of Communication* 35, no. 2 (2020): 126-139. https://doi.org/10.1177/0267323119894489

Anspach, Nicolas M. "The new personal influence: How our Facebook friends influence the news we read." *Political Communication* 34, no. 4 (2017): 590-606.

Anspach, Nicolas M., and Taylor N. Carlson. "What to believe? Social media commentary and belief in misinformation." *Political Behavior* 42, no. 3 (2020): 697-718.

Balfour, Virginia H. "Likes, comments, action! An examination of the Facebook audience engagement strategies used by strategic impact documentary." *Media International Australia* 176, no. 1 (2020): 34-51.

Bastick, Zach. "Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation." Computers in Human Behavior, 116 (in press 2021): 1-12.

Bechmann, Anja, and Kristoffer L. Nielbo. "Are we exposed to the same "News" in the news feed?". *Digital Journalism* 6, no. 8 (2018): 990-1002.

Bernal, Paul. "Facebook: Why Facebook makes the fake news problem inevitable." *Northern Ireland Legal Quarterly* 69 (2018): 513-530.

Beyens, Ine, Eline Frison, and Steven Eggermont. "'I don't want to miss a thing': Adolescents' fear of missing out and its relationship to adolescents' social needs, Facebook use, and Facebook related stress." *Computers in Human Behavior* 64 (2016): 1-8.

Binford, Matthew, T., Bartosz W. Wojdynski, Yen-I Lee, Shuoya Sun, and Andrea Briscoe. "Invisible transparency: Visual attention to disclosures and source recognition in Facebook political advertising." *Journal of Information Technology & Politics* 18, no.1 (2021): 70-83.

Błachnio, Agata, and Aneta Przepiórka. "Facebook intrusion, fear of missing out, narcissism, and life satisfaction: A cross-sectional study." *Psychiatry Research* 259 (2018): 514-19.

Blaine, Timothy, and Pascal Boyer. "Origins of sinister rumors: A preference for threat-related material in the supply and demand of information." *Evolution and Human Behavior* 39, no. 1 (2018): 67-75.

Boerman, Sophie C., Lotte M. Willemsen, and Eva P. Van Der Aa. ""This post is sponsored": Effects of sponsorship disclosure on persuasion knowledge and electronic word of mouth in the context of Facebook." *Journal of Interactive Marketing* 38 (2017): 82-92.

Bond, Sherman, and Allen, Bobby, "48 AGs, FTC Sue Facebook, Alleging Illegal Power Grabs to 'Neutralise' Rivals." *NPR*, December 9, 2020.

Borchers, Callum. "Facebook invites you to live in a bubble where you are always right." *The Washington Post*, January 14, 2018.

https://www.washingtonpost.com/news/the-fix/wp/2018/01/14/facebook-invites-you-to-live-in-a-bubble-where-you-are-always-right/.html

Borgman, Christine L. *Big data, little data, no data: Scholarship in the networked world*. Cambridge, Mass: MIT Press, 2015.

Bradford, Ben, Florian Grisel, Tracey L. Meares, Emily Owens, Baron L. Pineda, Jacob N. Shapiro, Tom R. Tyler, and Danieli E. Peterman. "Report of the Facebook Data Transparency Advisory Group." *The Justice Collaboratory: Yale Law School*: 1-43.

Bruns, Axel. "The Future is User-Led: The Path Towards Widespread Produsage." *Proceedings of Perth DAC 2007: The 7th International Digital Arts and Culture Conference, edited by* A. Hutchinson: 68-77. Curtin University of Technology, Western Australia, Perth, 2007.

Cabañas, González, Cuevas, Àngel, Arrate, Aritz, and Cuevas, Rubén, Communications of the ACM, 64, no. 1 (January 2021): 62-69.

Cain, Áine. "The 11 coolest perks at Facebook, the best place to work in America." *Business Insider*, December 7, 2017.

Cain, Áine. "What it is REALLY like to work at Facebook," *Business Insider*, December 7, 2017.

Carlson, Nicholas. "At Last – the Full Story of How Facebook Was Founded," *Tech Insider*, March 5, 2010.

Carmi, Elinor. "Rhythmedia: A study of Facebook immune system." *Theory, Culture and Society* 37, no. 5 (2020): 119-138.

Carnahan, Daniel. "Facebook has made AR ads available to all marketers through its ad manager." *Insider*, December 13, 2019.

Cetina Presuel, Rodrigo, and José Martínez Sierra. "Algorithms and the News: Social Media Platforms as News Publishers and Distributors." *Revista De Comunicación* 18, no. 2 (2019): 261-285.

Chander, Anupam, and Vivek Krishnamurthy. "The myth of platform neutrality." *2 Georgetown Law Technology Review 400*, (2018): 400-416.

Cheung, Christy, Pui-Yee Chiu, and Matthew K.O. Lee. "Online social networks: Why do students use Facebook?" *Computers in Human Behavior* 27 (2011): 1337-1343.

Chitra, Uthsav, and Christopher Musco. "Analyzing the Impact of Filter Bubbles on Social Network Polarization." WSDM '20 Proceedings of the 13th International Conference on Web Search and Data Mining. (January, 2020): 115-123. https://doi.org/10.1145/3336191.3371825

Choraś, Michał, Konstantinos Demestichas, Agata Giełczyk, Álvaro Herrero, Paweł Ksieniewicz, Konstantina Remoundou, Daniel Urda, Michał Woźniak. "Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study." *Applied Soft Computing Journal*, 101 (in press 2021).

Chou, Hui-Tzu (Grace), and Nicholas Edge. "'They are happier and having better lives than I am': The impact of using Facebook on perceptions of others' lives." *Cyberpsychology, Behavior & Social Networking* 15, no. 2 (2012): 117-121.

Chowdhry, Amit. "Facebook Launches A New Tool That Combats Fake News." *Forbes*, March 5, 2017.

Christofides, Emily, Amy Muise, and Serge Desmarais. "Information disclosure and control on Facebook: Are they two sides of the same coin or two different processes?" *CyberPsychology and Behavior* 12, no. 3 (2009): 341-345.

Chung, Myojung, and Nuri Kim. "When I learn the news is false: How fact-checking information stems the spread of fake news via third-person perception." *Human Communication Research*. 47, no. 1 (2021): 1-24. https://doi.org/10.1093/hcr/hqaa010

Cialdini, Robert, B. *Influence: The psychology of persuasion*. New York: Harper Collins, 1984.

Cialdini, Robert, B. *Pre-suasion: A revolutionary way to influence and persuade.* New York: Random House, 2016.

Cinelli, Matteo, Mauro Conti, Livio Finos, Francesco Grisolia, Petra K Novak, Antonio Peruzzi, Maurizio Tesconi, Fabiana Zollo and Walter Quattrociocchi. "(Mis)Information Operations." *Journal of Information Warfare* 18, no. 3 (2019): 83-98.

Clayton, Katherine, Spencer Blair, Jonathan A. Busam, Samuel Forstner, John Glance, Guy Green, Anna Kawata, *et al.* "Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media." *Political Behavior* 42, no. 4 (2020): 1073-1095.

Clegg, Nick. "You and the Algorithm: It Takes Two to Tango." *Nick Clegg*, March 31, 2020.

Clegg, Nick. "Welcoming the Oversight Board." *Facebook*, May 6, 2020. https://about.fb.com/news/2020/05/welcoming-the-oversight-board/

Coaston, Jane. "The Facebook free speech battle, explained." *Vox*, 14 May, 2019.

Constine, Josh. "A flaw-by flaw guide to Facebook's new GDPR privacy changes." *Tech Crunch*, April 18, 2018.

Cuevas, Angel, et al.. "Does Facebook Use Sensitive Data for Advertising Purposes? Worldwide Analysis and GDPR Impact." *Communications of the ACM* 64. No 1 (2021): 64. https://arxiv.org/abs/1907.10672v1

Cugelman, Brian, Mike Thelwall, and Phil Dawes. "Website credibility, active trust and behavioural intent." Paper presented at the *International Conference on Persuasive Technology*, Berlin, Heidelberg, 2008.

Cyr, Dianne, Milena Head, Eric Lim, and Agnis Stibe. "Using the elaboration likelihood model to examine online persuasion through website design." *Information & Management* 55, no. 7 (2018): 807-821.

De keersmaecker, Jonas, and Arne Roets. "'Fake news': Incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions." *Intelligence* 65 (2017): 107-110.

Del Vicario, Michela, Alessandro Bessi, Fabiana Zollo, Fabio Petronic, Antonio Scala, Guido Caldarelli, Eugene Stanley , and Walter Quattrociocchi. "The spreading of misinformation online." *PNAS* 113, no. 3 (2016): 554-559.

DeMarzo, Peter M., Dimitri Vayanos, and Jeffrey Zwiebel. "Persuasion bias, social influence, and unidimensional opinions." *The Quarterly Journal of Economics* 118, no. 3 (2003): 909-968.

DeVito, Michael A. "From Editors to Algorithms." *Digital Journalism* 5, no. 6 (2017): 753-773.

De la Merced, Michael J. "Facebook Closes at $38.23, Nearly Flat on Day." *New York Times*, May 18, 2012.

Diakopoulos, Nicholas. "Algorithmic accountability reporting: on the investigation of black boxes." *Columbia Journalism School. Tow Center for Digital Journalism*, 2014: 1-37. Algorithmic Accountability Reporting: On the Investigation of Black Boxes | Tow Center (columbia.edu)

Dias, Nicholas, and Amy Sippitt. "Researching Fact Checking: Present Limitations and Future Opportunities." *The Political Quarterly*, 91, No. 3 (2020): 605-613.

Di Pietro, Loredana, and Eleonora Pantano. "An empirical investigation of social network influence on consumer purchasing decision: The case of Facebook." *Journal of Direct, Data and Digital Marketing Practice* 14, no. 1 (2012): 18-29.

Dogruer, Nazan, Ipek Menevis, and Ramadan Eyyam. "What is the motivation for using Facebook?" *Procedia Social and Behavioral Sciences* 15 (2011): 2642-2646.

Doody, Amy. "Can anyone challenge Facebook's dominance in social?" *The Irish Times*, February 7, 2020. Can anyone challenge Facebook's dominance in social? (irishtimes.com)

Dubov, Alex. "Ethical persuasion: The rhetoric of communication in critical care." *Journal of Evaluation in Clinical Practice* 21, no. 3 (2015): 496-502.

Effron, Daniel A., and Medha Raj. "Misinformation and morality: Encountering fake-news headlines makes them seem less unethical to publish and share." *Psychological Science* 31, no. 1 (2020): 75-87.

Enke, Benjamin, and Florian Zimmermann. "Correlation neglect in belief formation". *The Review of Economic Studies*, 86, no. 1 (2019): 313-332.

Erisen, Cengiz, David P. Redlawsk, and Elif Erisen. "Complex thinking as a result of incongruent information exposure." *American Politics Research* 46, no. 2 (2017): 217-245.

https://doi.org/10.1177/1532673X17725864

Facebook, Inc. "Fact Checking on Facebook." *Facebook for Business*: https://www.facebook.com/business/help/2593586717571940

Fazio, Lisa K., Nadia M Brashier, B Keith Payne, and Elizabeth J Marsh. "Knowledge does not protect against illusory truth." *Journal of Experimental Psychology: General* 144, no. 5 (2015): 993-1002.

Federal Trade Commission, (US). "FTC Imposes a $5 Billion Penalty and Sweeping New Privacy Restrictions on Facebook." *Federal Trade Commission, press releases*, July 24, 2019.

Federal Trade Commission, (US). "Dissenting Statement of Commissioner Rohit Chopra." *United States of America Federal Trade Commission*, July 24, 2019.

Federal Trade Commission, (US). "Dissenting Statement of Commissioner Rebecca Kelly Slaughter." *United States of America Federal Trade Commission*, July 24, 2019.

Federal Trade Commission, (US). "FTC Sues Facebook for Illegal Monopolisation." *Federal Trade Commission press releases*, December 9, 2020.

Feezell, Jessica T. "Agenda setting through social media: The importance of incidental news exposure and social filtering in the digital era." *Political Research Quarterly* 71, no. 2 (2018): 482-494.

Feloni, Richard. "Mark Zuckerberg shares the hiring rule he says separates good companies from great ones." *Business Insider*, June 13, 2017.

Fisher, Christine. "Wikipedia co-founder wants to give you an alternative to Facebook and Twitter." *Engadget*, November 14, 2019.

Fletcher, Richard, Alessio Cornia, Lucas Graves, and Rasmus Kleis Neilsen. *"Measuring the Reach of "Fake News" and Online Disinformation in Europe."* Oxford: University of Oxford, 2018.

Flyverbom, Mikkel. "Transparency: Mediation and the management of visibilities." *International Journal of Communication* 10 (2016): 110-122.

Forbes. "Open Facebook." September 11, 2006.

Fogg, Brian J., and Dean Eckles. "The behavior chain for online participation: How successful Web services structure persuasion." Paper presented at the *International Conference on Persuasive Technology*, Berlin, Heidelberg, 2007.

Fogg, Brian J., and Daisuke Iizawa. "Online persuasion in Facebook and Mixi: A cross-cultural comparison." Paper presented at the *International Conference on Persuasive Technology*, Berlin, Heidelberg, 2008.

Fox, Gary. "Facebook Business Model: How Does Facebook Make Money." *Kerry Fox*, March 8, 2020.

Freeze, Melanie, Mary Baumgartner, Peter Bruno, Jacob R. Gunderson, Joshua Olin, Morgan Quinn Ross, and Justine Szafran. "Fake claims of fake news: Political misinformation, warnings, and the tainted truth effect." *Political Behavior* (in press, 2020): 1-33.

Frenkel, Sheera, et al.. "Delay, Deny, Deflect: How Facebook Leaders Fought Through Crisis." *New York Times*, November 15, 2018.

Fridkin, Kim, Patrick J. Kenney, and Amanda Wintersieck. "Liar, liar, pants on fire: How fact-checking influences citizens' reactions to negative advertising." *Political Communication* 32, (2015): 127-151. doi:10.1080/10584609.2014.914613

Frier, Sarah and Grant, Nico. "Instagram Brings in More Than a Quarter of Facebook Sales." Bloomberg, February 5, 2020.

Fumagalli, Andrea, Stefano Lucarelli, Elena Musolino, and Giulia Rocchi. "Digital labour in the platform economy: The case of Facebook." *Sustainability* 10, no. 6 (2018): 1757, 1-16.

Garrett, R. Kelly, Brain F. Weeks, B.E., and Rachel L. Neo. "Driving a wedge between evidence and beliefs: How online ideological news exposure promotes political misperceptions." *Journal of Computer-Mediated Communication* 21 (2016): 331-348.

Geschke, Daniel, Jan Lorenz, and Peter Holtz. "The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers." *British Journal of Social Psychology* 58, no. 1 (2018): 129-149.

Ghaffray, Shirin, and Rebecca Heilweil. "Why Facebook is 'the front line' in fighting hate today." *Recode*, 15 July, 2020. Retrieved 2 March, 2021 from:

https://www.vox.com/recode/2020/7/15/21325728/facebook-stop-hate-for-profit-campaign-jonathan-greenblatt-anti-defamation-league

Giachanou, Anastasia, Esteban Ríssola, Bilal Ghanem, Fabio Crestani, and Paolo Rosso. "The Role of Personality and Linguistic Patterns in Discriminating Between Fake News Spreaders and Fact Checkers." *International Conference on Applications of Natural Language Systems*. NLDB 2020: Natural Language Processing and Information Systems (2020): 181-192.

Gillett, Rachel. "7 reasons Facebook is the best place to work in America." *Business Insider*, December 7, 2017.

Golbeck, Jennifer, Matthew Mauriello, Brooke Auxier, Keval H. Bhanushali, Christopher Bonk, Mohamed Amine Bouzaghrane, et al.. "Fake News vs Satire: A Dataset and Analysis." WebSci'18, Amsterdam, Netherlands: Proceedings of the 10th ACM Conference on Web Science (May 27-30, 2018): 17-21.

Goldman, Eric. "Want to Kill Facebook and Google? Preserving Section 230 Is Your Best Hope." Balkinization: *New Controversies in Intermediary Liability Law,* (2019): 1-5. Want to Kill Facebook and Google? Preserving Section 230 Is Your Best Hope by Eric Goldman :: SSRN

Goodwin, Bill, and Boros, Crina. "Facebook undermined rivals in bid to dominate global messaging." *Computer Weekly*, November 26, 2019.

Gravanis, Georgios, Athena Vakali, Konstantinos Diamantaras, and Panagiotis Karadais. "Behind the cues: A benchmarking study for fake news detection." *Expert Systems and Applications* 128 (2019): 201-213.

Greco, Elizabeth. "More Americans are turning to multiple social media sites for news." *Pew Research Center Factank*, November 2, 2017.

Grimmelmann, James. "Saving Facebook." *Iowa Law Review* 94 (2009): 1137-1151.

Grosser, Benjamin. "What do metrics want? How quantification prescribes social interaction on Facebook." *Computational Culture*, no. 4 (2014): 1-41.

Gorwa, Robert, and Timothy G. Ash. "Democratic transparency in the platform society." In *Social media and democracy: The state of the field, prospects for reform*, edited by Nate Persily, and Josh A. Tucker, 286-312. Cambridge: Cambridge university Press, 2020.

Guadagno, Rosanna, E., and Karen Guttieri. "Fake news and information warfare: An examination of the political and psychological processes from the digital sphere to the real world." In *Handbook of research on deception, fake news, and misinformation online*, edited by Innocent E. Chiluwa, and Sergei A. Samoilenko, 167-191. Hershey, PA.: IGI Global, 2019.

Haglin, Kathryn. "The limitations of the backfire effect." Research and Politics, July-September, 2017: 1-5. The limitations of the backfire effect (sagepub.com)

Harris, Lisa, and Charles Dennis. "Engaging customers on facebook: Challenges for e-retailers." *Journal of Consumer Behaviour* 10, no. 6 (2011): 338-346.

Heath, Nick. "Trapped in a filter bubble: How Google and Facebook only feed you what you want to hear." *ZDnet*, July 18, 2011.

https://www.zdnet.com/article/trapped-in-a-filter-bubble-how-google-and-facebook-only-feed-you-what-you-want-to-hear/

Helberger, Natali. "The Political Power of Platforms: How Current Attempts to Regulate Misinformation Amplify Opinion Power." *Digital Journalism* 8, no. 6 (2020): 842-854.

Henderson, Samantha, and Michael Gilding. "'I've never clicked this much with anyone in my life': Trust and hyperpersonal communication in online friendships." *New Media & Society* 6, no. 4 (2004): 487-506.

Hew, Khe Foon. "Students' and teachers' use of Facebook." *Computers in Human Behavior* 27, no. 2 (2011): 662-76.

Horne Benjamin D., Sibel Adah. "This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News." The Workshops of the Eleventh International AAAI Conference on Web and Social Media AAAI Technical Report, News and Public Opinion (2017): 1-9.

Horwitz, Jeff, and Deepa Seetharaman. "Facebook Executives Shut Down Efforts to Make the Site Less Divisive." *Wall Street Journal*, May 26, 2020. Retrieved 18 March, 2021 from Facebook Executives Shut Down Efforts to Make the Site Less Divisive - WSJ

House of Commons Digital, Culture, Media and Sports Committee. "*Disinformation and 'Fake News': Final Report, Eighth Report of Session 2017-19*". London: United Kingdom House of Commons, 2019.

House of Representatives, (US). "Investigation of Competition in Digital Markets: Majority Staff Report and Recommendations." *Subcommittee on Anti-Trust, Commercial and Administrative Law of the Committee on the Judiciary*, (Washington: US House of Representatives, 2020).

Houser, Kimberly A. and Voss, W. Gregory. "GDPR: the end of Google and Facebook or a New Paradigm in Data Privacy?" *Richmond Journal of Law and Technology* 25, no. 1 (2018).

Huang, Hanjing, Hengameh Akbaria, Nina Alef, Phairoj Liukitithara, Monica Marazzi, Bastian Verhaelen, Gina Chi-Lan Yang, and Pei-Luen Patrick Rau. "International Users' Experience of Social Media: A Comparison Between Facebook and WeChat." *10th International Conference*, CCD 2018. Held as Part of HCI International 2018 Las Vegas, NV, USA, July 15–20, 2018, Proceedings, Part I, Pei-Luen Patric Rau (ed.): 341-349.

Ingram, David. "Exclusive: Facebook to put 1.5 billion users out of reach of new EU privacy law." *Reuters*, April 19, 2018.

Jackson, Dominique. "11 Facebook Metrics Every Brand Needs to Track." *Spout Blog*, August 7, 2020.

Kaplan, Katherine. "Facemash Creator Survives Ad Board." *The Harvard Crimson*, November 19, 2003.

Krombholz, Katharina, et al.. "Fake identities in social media." *Journal of Service Science*, no 4, 2012, 175-212.

Larson, Eric. "Twitter, Facebook Win Appeal in Anticonservative Bias Suit." *Bloomberg*, 28 May, 2020.

Losifidis, Petros, and Nicholas Nicoli. "The battle to end fake news: A qualitative content analysis of Facebook announcements on how it combats disinformation." *The International Communication Gazette* 82, no. 1 (2020): 60-81.

Jacobs, Naomi. "Two ethical concerns about the use of persuasive technology for vulnerable people." *Bioethics* 34, no. 5 (2020): 519-26.

Jiang, Shan, and Christo Wilson. "Linguistic Signals under Misinformation and Fact-Checking: Evidence from User Comments on Social Media." *Proceedings of the ACM on Human-Computer Interaction 2*, No. CSCW, Article 82 (November, 2018): 1-23.

Jones-Jang, Mo, Tara Mortensen, and Jing-Jing Lui. "Does Media Literacy Help Identification of Fake News? Information Literacy Helps, but Other Literacies Don't." *American Behavioral Scientist* 65, no. 2 (2019): 1-18.

Jouhki, Jukka, Epp Lauk, Maija Penttinen, Niina Sormanen, and Turo Uskali. "Facebook's Emotional Contagion Experiment as a Challenge to Research Ethics." *Media and Communication* 4, no. 4, 2016: 75-85.

Jowitt, Tom. "Zuckerberg Proposes 'Thoughtful Reform' Of Section 230 Liability." *Silicon*, March 25, 2021.

Just, Natascha, and Michael Latzer. "Governance by algorithms: reality construction by algorithmic selection on the Internet." *Media, Culture & Society* 39, no. 2 (2017): 238-258.

Karnowski, Veronika, Anna S. Kümpel, Larissa Leonhard, and Dominik J. Leiner. "From incidental news exposure to news engagement. How perceptions of the news post and news usage patterns influence engagement with news articles encountered on Facebook." *Computers in Human Behavior* 76 (2017): 42-50.

Kelly, Dimitri. "Evaluating the news: (Mis)perceptions of objectivity and credibility." *Political Behavior* 41, no. 2 (2019): 445-471.

Kelly, Heather. "Mark Zuckerberg explains why he just changed Facebook's mission." *CNN Business*, 22 June, 2017. Retrieved 1 March, 2021 from: https://money.cnn.com/2017/06/22/technology/facebook-zuckerberg-interview/index.html?iid=EL

Kelly, Makena. "Biden's Commerce nominee backs changes to Section 230." *The Verge*, January 26, 2021.

Kim, Antino, Patricia L. Moravec, and Alan R. Dennis. "Combating Fake News on Social Media with Source Ratings: The Effects of User and Expert Reputation Ratings." *Journal of Management Information Systems* 36, no. 3, (2019): 931-968.

Kjelsaas, Kathleen, Naomi Mecklenburgh, Courtney von Hippel, and Loren Brener. "For better or worse: Facebook use and wellbeing among community health care clients." *The Journal of Nervous and Mental Disease* 207, no. 8 (2019): 659-667.

Koebler, Jason, and Cox Joseph. "Internal documents show Facebook has never deserved our trust or our data." *Vice*, https://www.usatoday.com/story/tech/columnist/komando/2016/03/04/dont-click-like-facebook-again-until-you-read/81264440/

Koschate, Miriam, Elahe Naserian, Luke Dickens, Avelie Stuart, Alessandra Russo, and Mark Levine. "ASIA: Automated Social Identity Assessment using linguistic style." *Behavior Research Methods* (in press, 2021) https://doi.org/10.3758/s13428-020-01511-3

Kozinets, Robert, Dolbec, Pierre-Yann and Earley, Amanda. "Netnographic Analysis: Understanding Culture through Social Media Data." In *Sage Handbook of Qualitative Data Analysis*, ed. Uwe Flick (London: Sage, 2014), 262-272.

Kramer, Adam D.I., Jamie E. Guillory, and Jeffrey T. Hancock. "Experimental evidence of massive-scale emotional contagion through social networks." *Proceedings of the National Academy of Sciences* 111, no. 24 (2014): 8788-8790.

Krishan, Nihal. "Top House conservative introduces bill to overhaul Section 230 to curb child pornography." *Washington Examiner*, March 18, 2021.

Krombholz, Katharina, Dieter Merkl, and Edgar Weippl. "Fake identities in social media: A case study on the sustainability of the Facebook business model." *Journal of Service Science Research* 4, no. 2 (2012): 175-212.

Kross, Ethan, Philippe Verduyn, Emre Demiralp, Jiyoung Park, David S. Lee, Natalie Lin, Holly Shablack, John Jonides, Oscar Ybarra. "Facebook use predicts declines in subjective well-being in young adults." *PLoS ONE* 8 no. 8 (2013): e69841. Facebook Use Predicts Declines in Subjective Well-Being in Young Adults (plos.org)

Kruikemeier, Sanne, Minem Sezgin, and Sophie C Boerman. "Political microtargeting: Relationship between personalized advertising on Facebook and voters' responses." *Cyberpsychology, Behavior, and Social Networking* 19, no. 6 (2016): 367-372.

Kwon, Ohbyung, and Yixing Wen. "An empirical study of the factors affecting social network service use." *Computers in Human Behavior* 26, no. 2 (2010): 254-63.

Lazer, David MJ, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, *et al.*. "The science of fake news." *Science* 359, no. 6380 (2018): 1094-1096.

Lee, David. "Tech forms hail 'progress' on blocking terror." *BBC News*, June 8, 2018.

Lee, David. "Key fact-checkers stop working with Facebook." *BBC News,* Feburary 2, 2019. Available online at: https://www.bbc.com/news/technology-47098021

Lee, Jong Gun, Panayotis Antoniadis, and Kavé Salamatian. "Faving reciprocity in content sharing communities: A comparative analysis of Flickr and Twitter." Paper presented at the *2010 International Conference on Advances in Social Networks Analysis and Mining, Odense*, Denmark, August 9-11, 2010.

Lee, Sun Kyong, Nathan J. Lindsey, and Kyun Soo Kim. "The effects of news consumption via social media and news information overload on perceptions of journalistic norms and practices." *Computers in Human Behavior* 75 (2017): 254-263.

Levi, Or, Pedram Hosseini, Mona Diab, and David Broniatowski. "Identifying Nuances in Fake News vs. Satire: Using Semantic and Linguistic Cues." *Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda*, 2019: 31-35. 10.18653/v1/D19-5004

Levin, Sam. "Facebook told advertisers it can identify teens feeling 'insecure' and worthless." *The Guardian*, May 2, 2017. Retrieved 18 March, 2021 from: Facebook told advertisers it can identify teens feeling 'insecure' and 'worthless' | Facebook | The Guardian

Levin, Sam. "Is Facebook a publisher? In public it says no, but in court it says yes." *The Guardian*, July 3, 2018.

Lewandowsky, Stephan, Ullrich K.H. Ecker, Colleen M Seifert, Norbert Schwarz, and John Cook. "Misinformation and its correction: Continued influence and successful debiasing." *Psychological Science in the Public Interest* 13, no. 3 (2012): 106-131.

Lilley, Stephen, S. Grodzinsky Frances, and Andra Gumbus. "Revealing the commercialized and compliant Facebook user." *Journal of Information, Communication and Ethics in Society* 10, no. 2 (2012): 82-92.

Lim Heejin, and David W. Schumann. "Employing a Dramaturgical Lens to the Interpretation of Brand Online Social Networking: Evidence of Augmented Self." *Qualitative Market Research: An International Journal* 22, no. 3 (2019): 278-300. Employing a dramaturgical lens to the interpretation of brand online social networking: Evidence of augmented self | Emerald Insight

Luhmann, Niklas. *Trust and Power*. New York: Wiley, 1979.

MacLellan, Lila. "There are only four types of Facebook users, researchers have found." Quartz, July 13, 2017. https://qz.com/1026914/the-four-types-of-facebook-users-relationship-builders-window-shoppers-town-criers-and-selfies/

Maher, David. "Why Zuckerberg's Privacy-Focused Vision for Facebook Is Not Entirely Pro-Privacy." *Fortune*, March 7, 2019.

Marin, Lavinia. "Sharing (mis)information on social networking sites. An exploration of the norms for distributing content authored by others." *Ethics and Information Technology* (in press, 2021).

Marret, Christophe. "The relative impact of the Facebook's algorithm on the creation of social bubbles." (2020) (PDF) The relative impact of the Facebook's algorithm on the creation of social bubbles (researchgate.net)

Matz, Sandra C., Michal Kosinski, Gideon Nave, and David J Stillwell. "Psychological targeting as an effective approach to digital mass persuasion." *Proceedings of the national academy of sciences* 114, no. 48 (2017): 12714-19.

Mazarr, Michael J., Ryan Bauer, Abigail Casey, Sarah Heintz, and Luke J. Matthews. "The Emerging Risk of Virtual Societal Warfare: Social Manipulation in a Changing Information Environment." *Santa Monica, CA: RAND Corporation*, 2019. https://www.rand.org/pubs/research_reports/RR2714.html

McEvoy, Jemima. "Study: Facebook Allows And Recommends White Supremacist, Anti-Semitic And QAnon Groups With Thousands Of Members." *Forbes,* August 4, 2020. Retrieved March 1, 2021 from:

https://www.forbes.com/sites/jemimamcevoy/2020/08/04/study-facebook-allows-and-recommends-white-supremacist-anti-semitic-and-qanon-groups-with-thousands-of-members/?sh=1d0b993e6bbd

McKeown, Caitlin. "Facebook, defamation, and terrorism: Who is responsible for dangerous posts on social media?" *Tulane Journal of International and Comparative Law* 26 (2017): 163-187.

McLachlan, Stacy. "27 Facebook demographics to inform your strategy in 2021."

https://blog.hootsuite.com/facebook-demographics/#:~:text=6.%20Facebook%20reaches%20the%20largest%20number%20of%20users%20aged%2013%2D17andtext=Instagram%20reach%20for%20the%2013,teens%20compared%20to%20last%20year

McNamee, Roger. *Zucked: Waking up to the Facebook catastrophe.* Penguin Books, 2020.

Meade, Amanda, Taylor Josh, and Hurst, Daniel. "Facebook reverses Australia news ban after government makes media code amendments." *The Guardian*, February 23, 2021.

Mintz, Anne P. *Web of deceit : misinformation and manipulation in the age of social media.* Medford, N.J.: CyberAge Books, 2012.

Mohseni, Sina, Eric D. Ragan, and Xia Hu. "Open Issues in Combating Fake News: Interpretability as an Opportunity." *Computer Science: Social and Information Networks*. (2019): 1-7 [1904.03016] Open Issues in Combating Fake News: Interpretability as an Opportunity (arxiv.org)

Moreno, Megan A, Rajitha Kota, Shari Schoohs, and Jennifer M Whitehill. "The Facebook influence model: A concept mapping approach." *Cyberpsychology, Behavior, and Social Networking* 16, no. 7 (2013): 504-511.

Nadkarni, Ashwini, and Stefan G Hofmann. "Why do people use Facebook?". *Personality and Individual Differences* 52, no. 3 (2012): 243-249.

New Zealand Ministry of Foreign Affairs and Trade. *Christchurch Call to eliminate terrorist & violent extremis content online.* Undated.

Neate, Rupert. "Over $119bn wiped off Facebook's market cap after growth shock." *The Guardian*, July 27, 2018.

Nelson-Field, Karen, Erica Riebe, and Kellie Newstead. "The emotions that drive viral video." *Australasian Marketing Journal* 21, no. 4 (2013): 205-211 The Emotions that Drive Viral Video (sagepub.com)

Nemr, Christina and Gangware, William. *Weapons of Mass Distraction: Foreign State-Sponsored Disinformation in the Digital Age,* (Plymouth, Minnesota: Park Advisors, 2019).

Newton, Casey. "The Trauma Floor: The secret lives of Facebook moderators in America." *The Verge*, February 25, 2019.

Nickerson, Raymond S. "Confirmation bias: A ubiquitous phenomenon in many guises." *Review of General Psychology* 2, no. 2 (1998): 175-220.

Nieminen, Sakari, and Lauri Rapeli. "Fighting Misperceptions and Doubting Journalists' Objectivity: A Review of Fact-checking Literature." *Political Studies Review* 17, no. 3 (2019): 296-309.

Ning Shen, Kathy, and Mohamed Khalifa. "Exploring multidimensional conceptualization of social presence in the context of online communities." *International Journal of Human−Computer Interaction* 24, no. 7 (2008): 722-748.

Nyhan, Brendan, and Jason Reifler. "When corrections fail: The persistence of political misperceptions." *Political Behavior* 32, no. 2 (2010): 303-330.

Nyhan, Brendan, and Jason Reifler. "Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information." *Vaccine* 33, no. 3 (2015): 459-464.

Oboler, Andre. "The rise and fall of a Facebook hate group." *First Monday 13,* no. 11 (2008): https://doi.org/10.5210/fm.v13i11.2254

Oeldorf-Hirsch, Anne, Mike Schmierbach, Alyssa Appelman, and Michael P. Boyle. "The Ineffectiveness of Fact-Checking Labels on News Memes and Articles." *Mass Communication and Society* 23, no. 5 (2020): 682-704. DOI: 10.1080/15205436.2020.1733613.

Orsoa, Daniele, Nicola Federicia, Roberto Copettic, Luigi Vetrugnoa, and Tiziana Bovea. "Infodemic and the spread of fake news in the COVID-19-era." *European Journal of Emergency Medicine* 27, no. 5 (2020): 327-328.

Oshikawa, Ray, Jing Qian, William Yang Wang. "A Survey on Natural Language Processing for Fake News Detection." *Proceedings of the 12th Conference on Language Resources and Evaluation* (LREC 2020), Marseille, May 11–16, 2020: 6086–6093.

Pantic, Igor. "Online Social Networking and Mental Health." *Cyberpsychology, Behavior, and Social Networking* 17, no. 10 (2014): 652-657.

Pariser, Eli. *The filter bubble: What the internet is hiding from you.* UK: Penguin, 2011.

Paul, Katie and Wolfe, Jan. "U.S. FTC probes Facebook's acquisition practices." *Reuters*, August 2, 2019.

Pennycook, Gordon, Tyrone D. Cannon, and David G. Rand. "Prior exposure increases perceived accuracy of fake news." *Journal of Experimental Psychology: General* 147, no. 12 (2018): 1865-1880.

Pennycook, Gordon, Ziv Epstein, Mohsen Mosleh, Antonio A. Arechar, Dean Eckles, and David G. Rand. "Shifting attention to accuracy can reduce misinformation online." *Nature* (2021, in press): 1-6.

Pennycook, Gordon, Jonathon McPhetres, Yunhao Zhang, Jacjson G. Lu, and David G. Rand. "Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy nudge intervention." *Psychological Science* (2021, in press).

Pennycook, Gordon, and David G. Rand. "Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning." *Cognition* 188 (2019): 39-50.

Petty, Richard E., and John T. Cacioppo. *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag, 1986.

Pourghomi, Pardis, Milan Dordevic, and Fadi Safieddine. "Facebook Fake Profile Identification: Technical and Ethical Considerations." *International Journal of Pervasive Computing and Communications* 16, no. 1 (2020): 1-9.

Press, Gil. "Why Facebook triumphed over all other social networks." *Forbes*, April 8, 2018. Why Facebook Triumphed Over All Other Social Networks (forbes.com)

Prier, Jarred. "Commanding the trend: Social media as information warfare." *Strategic Studies Quarterly* 11, no. 4 (2017): 50-85.

Purtill, James. "Facebook thinks it won the battle of the media bargaining code - but so does the government." *ABC Science*, February 28, 2020.

Putnam, Robert D. *Bowling alone: The collapse and revival of American community*. Simon and Schuster, 2000.

Rauniar, Rupak., Greg Rawski, Jei Yang, and Ben Johnson. "Technology acceptance model (TAM) and social media usage: an empirical study on Facebook." *Journal of Enterprise Information Management* 27, no. 1 (2014): 6-30.

Rauschnabel, Philipp A, Sandra Praxmarer, and Bjorn S Ivens. "Social media marketing: How design features influence interactions with brand postings on Facebook." In *Advances in Advertising Research (Vol. III)*, 153-161: Springer, 2012.

Richards, Imogen and Woods, Mark. "Legal and Security Frameworks for Responding to Online Violent Extremism." In *The Handbook of Collective Violence: Current Developments and Understanding,* (London and New York: Routledge, 2020).

Ritzer, George, and Nathan Jurgenson. "Production, consumption, presumption." *Journal of Consumer Culture* 10, no. 1 (2010): 13-36.

Robinson, Tom., Clark Callahan, Kristoff Boyle, Erica Rivera, and Janice K. Cho. "I ❤ FB: A Q-Methodology analysis of why people 'like' Facebook." *International Journal of Virtual Communities and Social Networking* 9, no. 2 (2017): 46-61.

Rodriguez, Salvador. "Inside Facebook's 'cult-like' workplace, where dissent is discouraged and employees pretend to be happy all the time." *CNBC*, January 8, 2019.

Ross, Lee, Mark R Lepper, and Michael Hubbard. "Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm." *Journal of Personality and Social Psychology* 32, no. 5 (1975): 880-892.

Ruchansky, Natali et al.. "CSI: A Hybrid Deep Model for Fake News Detection." *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, (2017).

Statt, Nick. "Mark Zuckerberg just unveiled Facebook's new mission statement." *The Verge*, June 22, 2017.

Schreiner, Melanie, Thomas Fischer, and Rene Riedl. "Impact of content characteristics and emotion on behavioral engagement in social media: Literature review and research agenda." *Electronic Commerce Research*. (2019, in press): 1-17. Impact of content characteristics and emotion on behavioral engagement in social media: literature review and research agenda (springer.com)

Seargeant, Philip, and Caroline Tagg. "Social media and the future of open debate: A user-oriented approach to Facebook's filter bubble conundrum." *Discourse, Context & Media* 27 (2019): 41-48.

Security and Exchange Commission (US). "Facebook to Pay $100 Million for Misleading Investors about the Risks It Faced from Misuse of User Data." July 25, 2019.

Security and Exchange Commission (US), Form 10-K. *Annual report pursuant to section 13 or 15(D) of the Securities Exchange Act of 1934. For the fiscal year ended December 31, 2020. Facebook, Inc.* Washington: United States Securities and Exchange Commission, 2020. https://www.sec.gov/Archives/edgar/data/1326801/000132680116000043/fb-12312015x10k.htm

Sehgal, Vibhor, Ankit Peshin, Sadia Afroz, and Hany Farid. "Mutual hyperlinking among misinformation peddlers." arXiv:2104.11694v1 [cs.SI] April 20, 2021.

Shariatmadar, David. "Could language be the key to detecting fake news?" *The Guardian*. September 2, 2019. Retrieved March 19, 2021 from: Could language be the key to detecting fake news? | Language | The Guardian

Sherman, Len. "Why Facebook Will Never Change Its Business Model." *Forbes*, April 16, 2018.

Sherman, Len. "Zuckerberg's Broken Promises Show Facebook Is Not Your Friend." *Forbes*, May 23, 2018.

Shen, Kathy N., and Mohamed Khalifa. "Exploring Multidimensional Conceptualization of Social Presence in the Context of Online Communities." *International Journal of Human-Computer Interaction* 24, no. 7 (2008): 722-748.

Shin, Jieun, Lian Jian, Kevin Driscoll, and François Bar. "Political rumoring on Twitter during the 2012 US presidential election: Rumor diffusion and correction." *New Media & Society* 19, no. 8 (2017): 1214-1235.

Shrum, LJ. *The psychology of entertainment media: Blurring the lines between entertainment and persuasion.* New York: Routledge, 2012.

Silow-Carroll, Andrew. "How Facebook monitors harmful content: 5 takeaways." *The Times of Israel*, December 29, 2018. Retrieved February 26, 2021 from: https://www.timesofisrael.com/how-facebook-monitors-harmful-content-5-takeaways/

Silverman, Craig and Singer-Vine, Jeremy. "Most Americans Who See Fake News Believe It, New Survey Says." *Buzzfeed News*, December 6, 2016.

Singer, Natasha. "What you don't know about how Facebook uses your data." *The New York Times*, April 11, 2019. https://www.nytimes.com/2018/04/11/technology/facebook-privacy-hearings.html

Social Media News. *Social Media Statistics Australia*. January, 2020. https://www.socialmedianews.com.au/social-media-statistics-australia-january-2020/

Solon, Olivia. "Facebook is hiring moderators. But is the job too gruesome to handle?" *The Guardian*, May 4, 2017.

Solon, Olivia. "Underpaid and overburdened: the life of a Facebook moderator." *The Guardian*, May 25, 2017.

Solon, Olivia. "Is Facebook replaceable? Tech investor launches bid to 'start the process'." *The Guardian*, April 25, 2018. Is Facebook replaceable? Tech investor launches bid to 'start the process' | Facebook | The Guardian

Spiliotopoulos, Tasos and Ian Oakley. "Understanding Motivations for Facebook Use: Usage Metrics, Network Structure, and Privacy." *CHI 2013*, Paris, France (April 27-May 2, 2013): 1-10.

Stein, Tao, Erdeng Cheng, and Karan Mangla. "Facebook immune system." EuroSys Social Network Systems (SNS), Salzburg (April 10, 2011): 1-8. facebook-immune.pdf (mit.edu)

Su, Chris Chao, and Ngai Keung Chan. "Predicting social capital on Facebook: The implications of use intensity, perceived content desirability, and Facebook-enabled communication practices." *Computers in Human Behavior* 72 (2017): 259-268.

Sümer, Omer, Efe Bozkir, Thomas Kübler, Sven Grüner, Sonja Utz, Enkelejda Kasneci. "Fake News Perception: An Eye Movement Dataset on the Perceived Believability of News Stories." *Data in Brief* (2021), doi: https://doi.org/10.1016/j.dib.2021.106909

Surma, Jerzy. "Social exchange in online social networks. The reciprocity phenomenon on Facebook." *Computer Communications* 73 (2016): 342-346.

Susser, Daniel, Beate Roessler, and Helen Nissenbaum. "Online manipulation: Hidden influences in a digital world." *Georgetown Law Technology Review* 4, no. 1 (2018): 1-45.

Taigman, Yaniv, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf. "DeepFace: Closing the Gap to Human-Level Performance in Face Verification." *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 2014, 1701-1708, doi: 10.1109/CVPR.2014.220.

Tajfel, Henri. "Social identity and intergroup behaviour." *Social Science Information* 13, no. 2 (1974): 65-93.

Taylor, Josh. "Facebook v Apple: the looming showdown over data tracking and privacy." *The Guardian*, February 14, 2021.

Thompson, Ben. "Facebook's privacy cake." *Stratechery*, March 7, 2019.

Thornhill, Calum, Quentin Meeus, Jeroen Peperkamp, and Bettina Berendt. "A Digital Nudge to Counter Confirmation Bias." *Frontiers in Big Data* 2, no. 11 (2019): 1-9.

Thorson, Kjerstin, Kelley Cotter, Mel Medeiros, and Chankyung Pak. "Algorithmic inference, political interest, and exposure to news and politics on Facebook, Information." *Information Communication & Society* 24, no. 2 (2021): 183-200.

Tillman, Maggie. "10 reasons why Facebook has thrived for 15 years." *Pocket-lint*, February 4, 2019. 10 reasons why Facebook has thrived for 15 years (pocket-lint.com)

Tong, Stephanie Tom, Brandon Van Der Heide, Lindsey Langwell, and Joseph B Walther. "Too much of a good thing? The relationship between number of friends and interpersonal impressions on Facebook." *Journal of Computer-mediated Communication* 13, no. 3 (2008): 531-549.

Torabi Asr, Fatemeh. "The language gives it away: How an algorithm can help us detect fake news." *The conversation,* August 5, 2019. The language gives it away: How an algorithm can help us detect fake news (theconversation.com)

Trautman, Lawrence. "Governance of the Facebook Privacy Crisis." *Pittsburgh Journal of Technology Law and Policy* 20 (2019-2020).

Turcotte, Jason, Chance York, Jacob Irving, Rosanne M Scholl, and Raymond J Pingree. "News recommendations from social media opinion leaders: Effects on media trust and information seeking." *Journal of Computer-Mediated Communication* 20, no. 5 (2015): 520-535.

Van Couvering, Elizabeth. "Faces and charts: platform strategies for visualising the audience, the case of Facebook, Information." *Communication & Society*, (2021): 1-18. DOI: 10.1080/1369118X.2021.1874034

Van Dijck, José. "Facebook as a tool for producing sociality and connectivity." *Television & New Media* 13, no. 2 (2012): 160-176.

Vedejová, Dáša, and Vladimíra Čavojová. "Confirmation bias in information search, interpretation, and memory recall: evidence from reasoning about four controversial topics." *Thinking & Reasoning* (in press, 2021): 1-29. DOI: 10.1080/13546783.2021.1891967

Vile, John R. "Incitement to imminent lawless action." https://mtsu.edu/first-amendment/article/970/incitement-to-imminent-lawless-action.

Volkova, Svitlana, Kyle Shaffer, Jin Yea Jang, and Nathan Hodas. "Separating Facts from Fiction: Linguistic Models to Classify Suspicious and Trusted News Posts on Twitter." *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics* (Short Papers), pages 647–653, Vancouver, Canada, (July 30 - August 4, 2017). 647-653.

Vosoughi, Soroush, Deb Roy, and Sinan Aral. "The spread of true and false news online." *Science* 359, no. 6380 (2018): 1146-1151.

Wagner, Ben, Krisztina Rozgonyi, Marie-Therese Sekwenz, Jennifer Cobbe, and Jatinder Singh. "Regulating transparency? Facebook, Twitter and the German Network Enforcement Act." *FAT* '20, Barcelona, Spain (January 27-30, 2020): 1-11.

Waldman, Ari E. "Manipulating trust on Facebook." *Loyola Consumer Law Review* 29, no. 1 (2016a): 175-198.

Waldman, Ari E. "Privacy, sharing, and trust: The Facebook study." *Case Western Reserve Law Review* 67, no. 1 (2016b): 193-234.

Walter, Nathan, Jonathan Cohen, R. Lance Holbert, and Yasmin Morag. "Fact-Checking: A Meta-Analysis of What Works and for Whom." Political Communication 37, no. 3 (2020): 350-375.

Wason, Peter C. "On the failure to eliminate hypotheses in a conceptual task." *Quarterly Journal of Experimental Psychology* 12, no. 3 (1960): 129-140.

Weedon, Jen, William Nuland, and Alex Stamos. "Information operations and Facebook." Facebook, Inc.. https://www.mm.dk/wp-content/uploads/2017/05/facebook-and-information-operations-v1.pdf.

White, Candace L., and Brandon Boatwright. "Social media ethics in the data economy: Issues of social responsibility for using Facebook for public relations." *Public Relations Review* 46, no. 5 (2020): 1-7.

Wilson, Robert E., Samuel D Gosling, and Lindsay T Graham. "A review of Facebook research in the social sciences." *Perspectives on Psychological Science* 7, no. 3 (2012): 203-20.

Wong, Julia Carrie. "Facebook's underclass: as staffers enjoying lavish perks, contractors barely get by." *The Guardian*, September 26, 2017.

Wong, Queenie and Van Boom, Daniel. "Facebook's News Ban in Australia: Everything You Should Know." *C/net*, February 23, 2021.

Xie, Chunhui, Jagannadha Sri Harsha Putrevu, and Chelsea Linder. "Family, Friends, and Cultural Connectedness: A Comparison Between WeChat and Facebook User Motivation, Experience and NPS Among Chinese People Living Overseas." *9th International Conference, CCD 2017*. Held as Part of HCI International 2017 Vancouver, BC, Canada, July 9–14, 2017, Proceedings. 369-384.

Yoon, Gunwoo, Cong Li, Yi (Grace) Ji, Michael North, Cheng Hong, and Jiangmeng Liu. "Attracting comments: Digital engagement metrics on Facebook and financial performance, *Journal of Advertising* 47, no. 1 (2018): 24-37.

Young, Dannagal G., Kathleen H. Jamieson, Shannon Poulsen, and Abigail Goldring. "Fact-Checking Effectiveness as a Function of Format and Tone: Evaluating FactCheck.org and FlackCheck.org." *Journalism & Mass Communication Quarterly* 95, no. 1 (2018): 49-75.

Zarsky, Tal Z. "Privacy and manipulation in the digital age." *Theoretical Inquiries in Law* 20, no. 1 (2019): 157-188.

Tufekci, Zeynep. "Zuckerberg's so-Called Shift toward Privacy." *New York Times*, March 7, 2019.

Zetlin, Minda. "Here's Why Facebook's Former Employees Describe the Company as Cult-Like." Inc, Undated.

Zhang, Xichen, Ali A. Ghorbani. "An overview of online fake news: Characterization, detection, and discussion." Information Processing and Management 57, no. 2 (2020): 1-26.

Zuboff, Shoshana. "Big other: Surveillance capitalism and the prospects of an information civilization." *Journal of Information Technology* 30, no. 1 (2015): 75-89.

Zuckerberg, Mark. "A Privacy Focused Vision for Social Networking." *New York Times*, March 6, 2019.

Zuckerberg, Mark. "Building Global Community." Facebook, May 6, 2021. https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10103508221158471

Appendices

1. List of Subsidiaries, Facebook, Inc.

1. Andale, LLC (Delaware)
2. Cassin Networks ApS (Denmark)
3. Facebook Holdings, LLC (Delaware)
4. Facebook Ireland Holdings Unlimited (Ireland)
5. Facebook Ireland Limited (Ireland)
6. Facebook Operations, LLC (Delaware)
7. Facebook Payments Inc. (Delaware)
8. Facebook Technologies, LLC (Delaware)
9. FCL Tech Limited (Ireland)
10. Goldframe LLC (Delaware)
11. Greater Kudu LLC (Delaware)
12. Instagram, LLC (Delaware)
13. KUSU PTE. LTD. (Singapore)
14. MALKOHA PTE LTD. (Singapore)
15. Morning Hornet LLC (Delaware)
16. Novi Financial, Inc. (Delaware)
17. Pinnacle Sweden AB (Sweden)
18. Raven Northbrook LLC (Delaware)
19. Runways Information Services Limited (Ireland)
20. Scout Development LLC (Delaware)
21. Siculus, Inc. (Delaware)
22. Sidecat LLC (Delaware)
23. Stadion LLC (Delaware)
24. Starbelt LLC (Delaware)
25. Vitesse, LLC (Delaware)
26. WhatsApp Inc. (Delaware)
27. Winner LLC (Delaware)
28. Woolhawk LLC (Delaware) [325]

---

[325] Form 10-K, Facebook, Inc., Exhibit 21.1.

2. List of companies acquired by Facebook, 2005-2021 (quoted in US$)

AboutFace: acquired August, 2005 – $200, 000

Parakeya: acquired on July 19, 2007

ConnectU: acquired on June 23, 2008 – $31 Million

FriendFeed: acquired on August 10, 2009 – $47.5 Million

Octazen: acquired on February 19, 2010

Divvyshot: acquired on March 2, 2010

Friendster: acquired on May 13, 2010 – $40 Million

ShareGrove: acquired on May 26, 2010

Nextstop: acquired on July 8, 2010 – $2.5 Million

Chai Labs: acquired on August 15, 2010 – $10 Million

Hot Potato: acquired on August 20, 2010 – $10 Million

Drop.io: acquired on October 29, 2010 – $10 Million

FB.com. domain name: acquired on November 15, 2010 – $8.5 Million

Rel8tion: acquired on January 25, 2011 – Undisclosed

BELUGA: acquired on March 2, 2011 – Undisclosed

Snaptu: acquired on March 20, 2011 – $70 Million

RecRec: acquired on March 24, 2011 – Undisclosed

DayTum: acquired on April 27, 2011

Sofa: acquired on June 9, 2011

MailRank: acquired on June 9, 2011

Push Pop Press: acquired on August 2, 2011 – Undisclosed

Friend.ly: acquired on October 10, 2011 – Undisclosed

Strobe: acquired on November 8, 2011 – Undisclosed

Gowall: acquired on December 2, 2011

Instagram: acquired on April 9, 2012 – $1 Billion

Tagtile: acquired on April 13, 2012 – Undisclosed

Glancee: acquired on May 5, 2012 – Undisclosed

Lightbox.com: acquired on May 15, 2012 – Undisclosed

Karma: acquired on May 21, 2012 – Undisclosed

Face.com: acquired on June 18, 2012 – $100 Million

Spool: acquired on July 14, 2012 – Undisclosed

Acrylic Software: acquired on July 20, 2012 – Undisclosed

Threadsy: acquired on August 24, 2012 – Undisclosed

Atlas: acquired on February 28, 2013 – Less than $100 Million

Osmeta: acquired March, 2013

Hot Studio: acquired on March 14, 2013

Spaceport: acquired on April 23, 2013

Parse: acquired on April 25, 2013 – $85 Million

Monoidics: acquired on July 18, 2013

Jibbigo: acquired on August 12, 2013

Onavo: acquired on October 13, 2013

SportStream: acquired on December 17, 2013

Little: **Eye Labs**: acquired on January 8, 2014 – $15 Million

Branch: acquired on January 13, 2014 – $15 Million

WhatsApp: acquired on February 19, 2014 – $19 Billion

Oculus VR: acquired on March 25, 2014 – $2 Billion

Ascenta: acquired on March 27, 2014 – $20 Million

Liverail: acquired on August 14, 2014 – $500 Million

ProtoGeoOy: acquired on April 24, 2014 – Undisclosed

Pryte: acquired June, 2014 – Undisclosed

PrivateCore: acquired on August 7, 2014 – Undisclosed

WaveGroupSound: acquired on August 26, 2014 – Undisclosed

Wit.ai: acquired on January 6, 2015 – Undisclosed

Quickfire: acquired on January 8, 2015 – Undisclosed

TheFind: acquired on March 14, 2015 – Undisclosed

Surreal Vision: acquired on May 26, 2015 – Undisclosed

Pebbles: acquired on July 16, 2015 – $60 Million

Masquerade: acquired on March 9, 2016 – Undisclosed

Two Big Ears: acquired on March 23, 2016 – Undisclosed

Nascent Objects: acquired on September 19, 2016 – Undisclosed

Infiniled: acquired on October 10, 2016 – Undisclosed

CrowdTangle: acquired on November 11, 2016 – Undisclosed

Faciometrics: acquired on November 16, 2016 – Undisclosed

Ozlo: acquired on July 31, 2017 – Undisclosed

Fayteq: acquired August 2017 – Undisclosed

tbh(app): acquired October, 2017 – Undisclosed

confirm.io: acquired on January 23, 2018 – Undisclosed

Bloomsbury AI: acquired July, 2018 – $30 Million

RedKix: acquired on July 26, 2018 – $100 Million

Vidpresso: acquired on August 13, 2018 – Undisclosed

Chainspace: acquired February, 2019 – Undisclosed

GROKSTYLE: acquired on February 8, 2019 – Undisclosed

Servicefriend: acquired September, 2019 – Undisclosed

CTRL-labs: acquired September, 2019 – Undisclosed, but reportedly between $500 Million and $1 Billion

Packagd: acquired September, 2019 – Undisclosed

Beat Games: acquired November, 2019 – Undisclosed

PlayGiga: acquired December, 2019 – $70 Million

Sanzaru Games: acquired February, 2020 – Undisclosed

Scape Technologies: acquired February, 2020 – Around $40 Million

Giphy: acquired on May 15, 2020 - $400 Million

Mapillary: acquired on June 18, 2020 – Undisclosed

Ready at Dawn: acquired on June 22, 2020 – Undisclosed

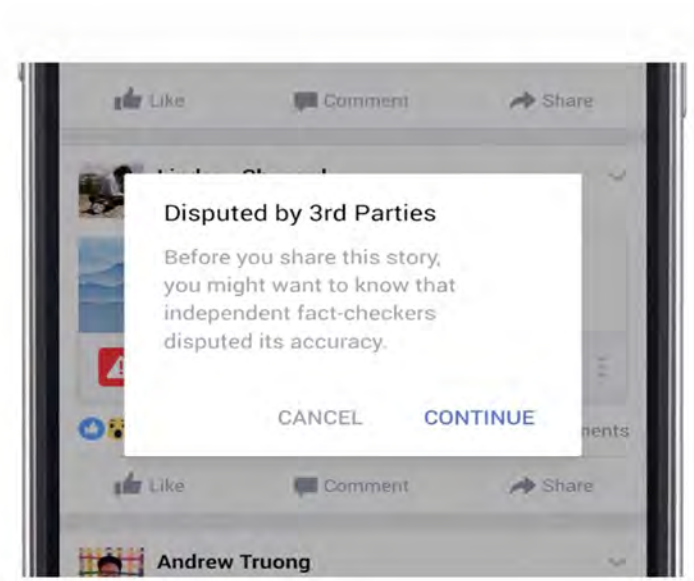**Lemnis Technologies:** acquired on September 18, 2020 – Undisclosed

Kustomer: acquired on November 30, 2020 – $1 Billion

Downpour Interactive: acquired on April 30, 2021 – Undisclosed [326]

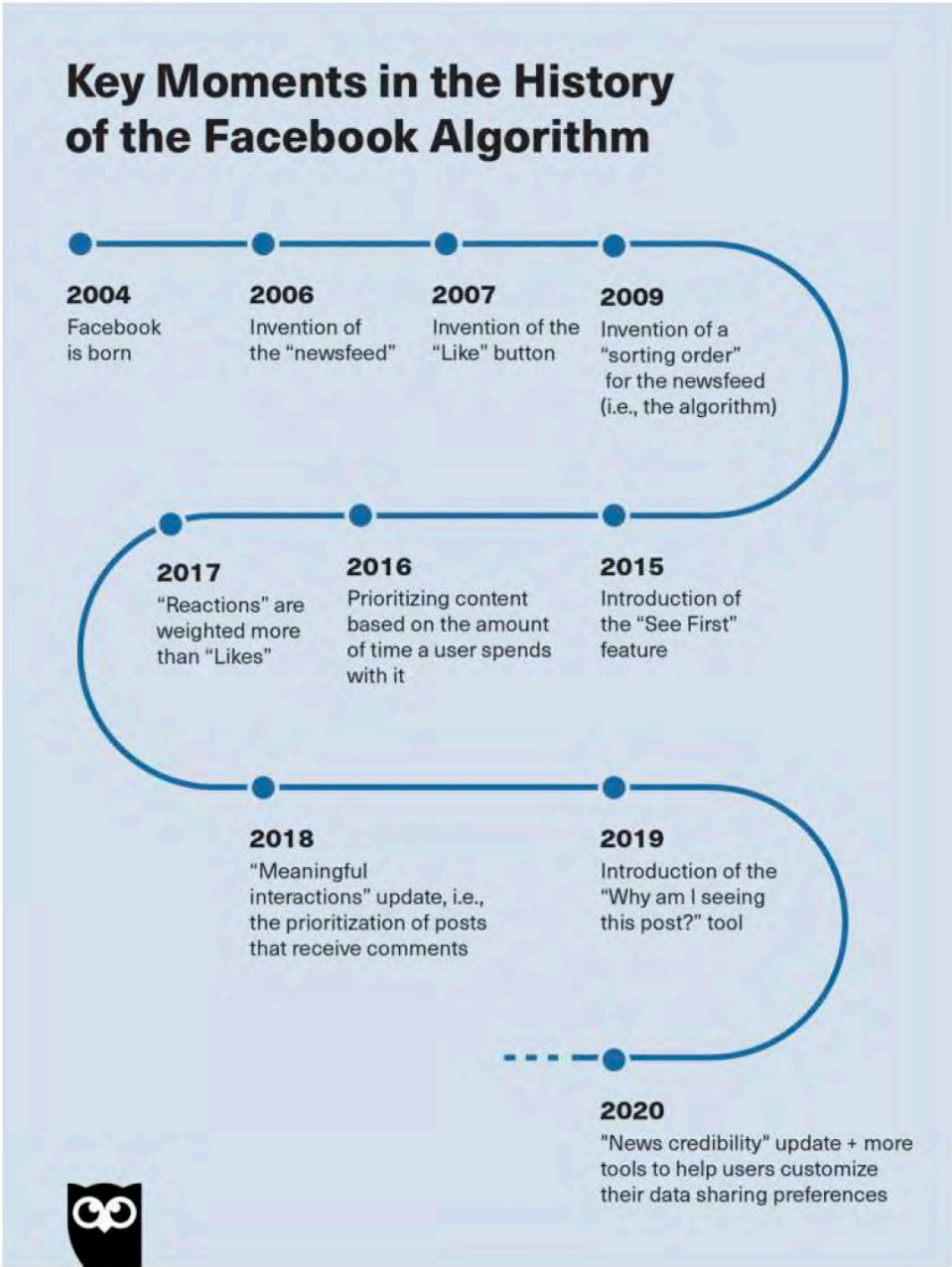[326] List of mergers and acquisitions by Facebook - Wikipedia

3. Examples of Facebook's tool for flagging 'fake' or at least disputed content [327]

[327] Amit Chowdhry, "Facebook Launches A New Tool That Combats Fake News," *Forbes*, March 5, 2017.

4. Key moments in the history of the Facebook algorithm

Key Moments in the History of the Facebook Algorithm

2004
Facebook is born

2006
Invention of the "newsfeed"

2007
Invention of the "Like" button

2009
Invention of a "sorting order" for the newsfeed (i.e., the algorithm)

2017
"Reactions" are weighted more than "Likes"

2016
Prioritizing content based on the amount of time a user spends with it

2015
Introduction of the "See First" feature

2018
"Meaningful interactions" update, i.e., the prioritization of posts that receive comments

2019
Introduction of the "Why am I seeing this post?" tool

2020
"News credibility" update + more tools to help users customize their data sharing preferences

5. Types of influence

Susser et al. locate different types of influence along a continuum with persuasion at one end, manipulation in the middle and coercion at the other end.[328] Persuasion is viewed as the least morally troubling form of influence, followed by manipulation and then coercion. Listed below are points of contrast between manipulation and persuasion:[329]

— Manipulation, like persuasion, is a type of influence that can be resisted. In the case of manipulation, however, it is not resisted, owing to the fact that the victim is not aware of the mechanism of influence. It therefore creates the illusion of free choice.

— Persuasion necessitates a trusting relationship based on mutual respect and interest; manipulation does not.

— Unlike persuasion, manipulation objectifies its victims and precludes the possibility of a relationship between manipulator and manipulated (at least in terms of an authentic relationship built on mutual respect/trust).

— Unlike manipulation, persuasion is usually viewed as a positive force, used to produce positive outcomes, at least insofar as it fosters free choice (even if one may disagree with the subject's choice).

— Persuasion does not involve deception but is, rather, the act of making the truth apparent. Although manipulation need not involve deception, typically it does.
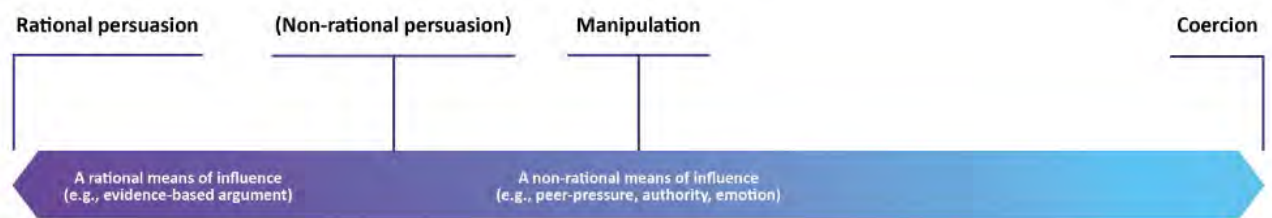

Figure 12: The continuum of influence
which includes non-rational persuasion positioned between rational persuasion and manipulation

Persuasion can also be thought of as involving an appeal to a subject's ability to reason – to their rational decision making – which contrasts with manipulation's covert attempt to circumvent or subvert this. In keeping with this view, Jacobs emphasises the *rational* nature of persuasion.[330] Rational persuasion influences by means of reason and argument (supported by evidence, where applicable). Thus, someone may come to believe something (be persuaded) because of the merit of the reasons another person advances.

Yet Jacobs also recognises that there is both a rational and non-rational means of persuasion (see Figure 12). Examples of non-rational persuasion include playing on someone's emotions (such as fear, anger or hatred, but also love and happiness), appealing to authority and peer pressure. Of crucial importance is that persuasion, whether rational or non-rational, if it is to count as a form of *persuasion*, must never significantly block or burden options. Moreover, even in the case of non-rational persuasion, unlike manipulation, a person must be aware of being intentionally influenced and understand the mechanisms of that influence (e.g., persuading someone to vote for a particular candidate by making salient their fear of crime and/or 'out of control' immigration). In

---

[328] Daniel Susser et al., "Online manipulation: Hidden influences in a digital world," *Georgetown Law Technology Review* 4, no. 1 (2018): 1-45.
[329] Alex Dubov, "Ethical persuasion: The rhetoric of communication in critical care," *Journal of Evaluation in Clinical Practice* 21, no. 3 (2015): 496-502.
[330] Naomi Jacobs, "Two ethical concerns about the use of persuasive technology for vulnerable people," *Bioethics* 34, no. 5 (2020): 519-26.

addition, a person is more likely to be persuaded (in either sense discussed) if they believe that what they are being influenced to do/feel/think is in their best interests and/or aligns with their personal goals and/or values.[331]

Further theoretical basis for non-rational persuasion

In 1986, Petty and Cacioppo proposed the *elaboration likelihood model* (ELM): a dual process model that posits central and peripheral routes to persuasion (see Figure 13). [332] By taking the peripheral route, one is persuaded not by the veracity of the claim and the quality of the supporting argument (i.e., the strength of evidence and reasoning), which is a characteristic of the central route, but instead by factors that act as decision heuristics,[333] such as emotional salience and/or rhetoric, the perceived *authority* of the speaker (a technique identified by Cialdini), the strength of agreement (i.e., consensus) and the similarity of these others to oneself (and one's social identity).
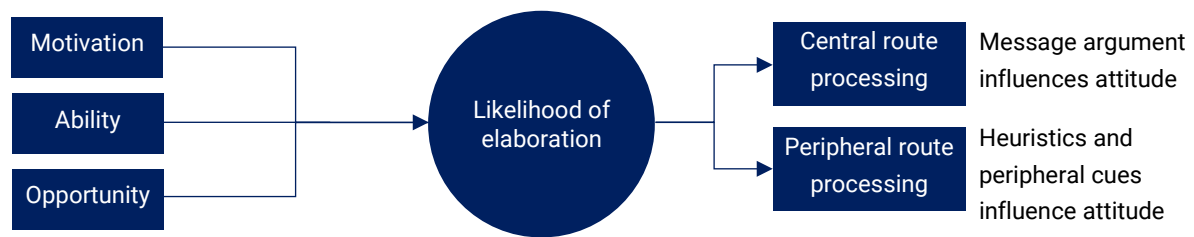

Figure 13: The elaboration likelihood model

Whether one employs the central or peripheral route is dependent on such things as the message's perceived relevance, one's prior knowledge of the topic (and therefore the degree of cognitive load required to process the information) and one's motivation. In short, it is dependent on motivation, ability and opportunity, and therefore one's state of *elaboration likelihood*.[334] Where the likelihood that one will engage in elaboration is low – that is, where it is unlikely that one would scrutinise the material presented for signs of reasoned argument and supporting evidence – then, as noted, it is more likely that one will employ decision heuristics and be persuaded by the material's aesthetic and/or emotional appeal, or the strength and nature of social presence or connectedness (i.e., trust or peer-pressure).[335] Lee et al., for example, found that the processing of text-based messages online is more likely to occur via the peripheral route.[336] In addition, Di Pietro and Pantano found, in support of the *Technology Acceptance Model* (TAM), that people were motivated to use new technology if it was thought to be useful, but also easy and fun to use, thereby reducing cognitive load and facilitating the use of decision heuristics.[337] One might conjecture that support for TAM correlates with what Moreno et al. called the *Facebook experience* (see Appendix 6),[338] whereby Facebook is experienced as a novel and, importantly, pleasant environment that enables users to connect easily with others or simply satisfy their curiosity about them, and/or express themselves (relatively) freely.

---

[331] Naomi Jacobs, "Two ethical concerns," 519-26.
[332] Petty, Richard E., and John T Cacioppo, Communication and persuasion: Central and peripheral routes to attitude change, New York: Springer-Verlag, 1986.
[333] Rosanna E. Guadagno, and Karen Guttieri, "Fake news and information warfare: An examination of the political and psychological processes from the digital sphere to the real world." In *Handbook of research on deception, fake news, and misinformation online*, edited by Innocent E. Chiluwa, and Sergei A. Samoilenko, 167-191. Hershey, PA.: IGI Global, 2019.
[334] Richard E. Petty, and John T. Cacioppo, *Communication and persuasion*, 1986.
[335] Dianne Cyr, Milena Head, Eric Lim, and Agnis Stibe, "Using the elaboration likelihood model to examine online persuasion through website design," *Information & Management* 55, no. 7 (2018): 807-821.
[336] Sun Kyong Lee , Nathan J. Lindsey, and Kyun Soo Kim, "The effects of news consumption via social media and news information overload on perceptions of journalistic norms and practices," *Computers in Human Behavior* 75 (2017): 254-263.
[337] Loredana Di Pietro, and Eleonora Pantano, "An empirical investigation of social network influence on consumer purchasing decision: The case of Facebook," *Journal of Direct, Data and Digital Marketing Practice* 14, no. 1 (2012): 18-29.
[338] Megan A. Moreno et al., "The Facebook influence model: A concept mapping approach," *Cyberpsychology, Behavior, and Social Networking* 16, no. 7 (2013): 504-511.

6. Facebook influence model clusters and example items

(taken from Moreno et al.) [339]

Domain	Cluster label	Example items within cluster
Connection	Connection to people	Allows people to constantly stay updated with others' lives; way to get to know acquaintances almost instantly; keep in touch with people you wouldn't call or text
	Far reaching	Ability to reach many people with one Web site; can reach anyone, young and old, rich and poor; bonding across cultures and distances
	Fast communication	Feel connected and in the loop constantly; puts everyone you know and what they're doing in one place; updates on people's lives faster than with a cell phone
	Business and promotion	Ability to plan influential events such as protests or sit-ins; statuses provide a way to blog instantly about events or political topics; every company uses it to promote business or provide deals
	Accessible and adaptable	Largest network in human history; easy to use and navigate; widely known and talked about
	Data and information	Huge database of information; compiled data from millions of individuals; news feature
Identification	Identity expression	Freedom to express things and let it be heard; present the best side of yourself; show off accomplishments to everyone you are friends with on Facebook, not just close friends
	Influence on identity	Provides others with pictures that can influence perceptions; display aspects of yourself that you wouldn't share in offline life (sexuality, substance use); wonder if you should be doing what you see everyone doing in pictures
Comparison	Curiosity about others	Can know what people are up to without asking them about it and without them knowing you know; creep culture/stalking; see who associates with whom with pictures and comments
	Facebook establishing social norms	Reinforces beliefs or opinions by seeing that others hold same beliefs or opinions; can see what is popular by observation; can follow norms
Facebook as an experience	Distractions	Procrastination; addictive; huge distraction
	Positive experiences	Facebook is referenced in daily life; provides entertainment at any time; status updates can promote a good mood
	Negative experiences	Changes the nature of communication from face to face to screen to screen; people willing to sacrifice privacy; inspires competition in people

---

[339] Megan A. Moreno et al., "The Facebook influence model," 504-511., p. 508.

7. Examples of Facebook advertising