Date: June 19, 2019

Does Digital Literacy Have a Role in Mitigating the Spread of Misinformation?

A Critical Analysis of Deepfake Videos and Their Comment Sections

Name: Dieuwertje Luitse
Student Number: 6558392

Pre-Master New Media &
Digital Culture, Utrecht University

Supervisor: Phd Alex P. Smit
Second Reader: Dr. Stefan Werning

Abstract

This study explores how digital literacy is thought of by commenters on YouTube in relation to the phenomenon of Deepfakes. The technology behind Deepfake videos is enabling the (algorithmic) creation of manipulated video footage of (high-profile) people that increasingly appears to be authentic and is feared to be become the next frontier in the battle against misinformation. Since becoming more visible as public problem after the 2016 U.S. presidential election, countries and organisations are doubling down on digital literacy as a solution to the problem of misinformation. Along with other authors, this study questions digital literacy as an individual solution to a societal problem. Because digital literacy focuses on knowledge and skills at the individual level, it is found to be too dependent on people's own assessments of media to effectively counter misinformation. Through a textual analysis of popular YouTube videos relating to Deepfakes and a critical discourse analysis of their comment sections, this study confirms that the skills associated with digital literacy are found to be effective at identifying the current generation of Deepfakes. Like many organisations, the majority of users who comment on the technology of Deepfakes express a concern about its harmful potential. While few of the over 3,000 comments analysed by this study concern digital literacy skills, most comments that put forward a solution focus on audio-visual validation skills that can falsify Deepfakes. Very few YouTube commenters suggest regulatory or technical solutions for combating the spread of Deepfakes. While this suggests that the public discourse on how to identify manipulated content is focussed on individual skills too, this study concludes that digital literacy should put more emphasis on increasing the public's knowledge of the role and responsibilities of the platforms that govern such content.

Keywords

Table of Content

1. Introduction: Deepfakes and the Call for Digital Literacy

In January 2019, a 'Deepfake' video of actors Jennifer Lawrence and Steve Buscemi started making waves as a demonstration of how far this new video-manipulation technology had come.[1] Deepfake is an AI-powered technology that allows users to "merge, combine, replace and superimpose images and video clips onto a video, creating a fake image or video that appears authentic."[2] While the Lawrence–Buscemi mashup was a humorous demonstration of the technology, many people worry about its harmful potential.[3,4] The makers of popular 'Deepfakes' are well aware of the manipulative power of the technology and demonstrate it making fabricated video statements from politicians such as former U.S. President Obama. This video explicitly aims to make the public aware of the new capabilities for manipulating video and emphasizes people's own responsibility in assessing the credibility of videos and images.[5]

Such calls can be connected to a broader call for media literacy and digital literacy. Media literacy concerns the creation and production of media messages, the ability to critically analyse these messages. Studies performed after the 2016 U.S. presidential election showed that many voters did not properly appreciate factual information, and that many voters feel this has aided the spread of misinformation. This plays to the idea that teaching individuals to become more media literate can help to fight misinformation.[6] Media literacy has "become a center of gravity" for countering misinformation, with many countries and organisations such as *The New York Times* and *The Washington Post* are investing in media-literacy education.[7] Meanwhile, companies like Google and Facebook announced investments in educational programmes for improving literacy skills amongst their users in a wider effort to fight the spread of misinformation.[8] Scholars have been critical of this

---

[1] Mikael Thalen (@MikaelThalen). "I've gone down a black hole of the latest DeepFakes and this mashup of Steve Buscemi and Jennifer Lawrence is a sight to behold." Tweet, January 29, 2019, https://twitter.com/MikaelThalen/status/ 1090349932266094593/video/1 9 (accessed February 13, 2019).

[2] Marie-Helen Maras and Alex Alexandrou. "Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos." *The International Journal of Evidence & Proof*, N.A. (2018): 1–8.

[3] VPRO Tegenlicht. "Deep Fake News." VPRO Tegenlicht, VPRO, 18 November 2018. https://www.vpro.nl/ programmas/tegenlicht/kijk/afleveringen/2018-2019/deep-fake-news.html (accessed March 10, 2019).

[4] Brian Klaas. "Deepfakes are coming. We're not ready." *The Washington Post*, May 14, 2019, *https://www.washingtonpost.com/opinions/2019/05/14/deepfakes-are-coming-were-not-ready/* (accessed May 18, 2019*).*

[5] BuzzFeedVideo. "You Won't Believe What Obama Says In This Video! 😉." Youtube video, 1:12, Apr 17, 2018. https://www.youtube.com/watch?v=cQ54GDm1eL0 (accessed March 10, 2019).

[6] Paul Mihailidis and Samantha Viotty. "Spreadable Spectacle in Digital Culture: Civic Expression, Fake News, and the Role of Media Literacies in "Post-Fact" Society." *American Behavioral Scientist*, 61.4 (2018): 441–454.

[7] Monica Bulger & Patrick Davison. *The Promises, Challenges, and Futures of Media Literacy*. New York: Data & Society Research Institute, 2018, 5. https://datasociety.net/pubs/oh/DataAndSociety_Media_Literacy_2018.pdf (accessed February 13, 2019).

[8] Dami Lee. "Facebook launches Digital Literacy Library to help young people use the internet responsibly." *The Verge,* August 2, 2019. https://www.theverge.com/2018/8/2/17643586/facebook-digital-literacy-library-educators-schools (accessed February 13, 2019).

development and have openly questioned whether digital literacy is the right solution to fight misinformation at large.[9, 10, 11]

1.1 Commenting on the Discourse on Digital Literacy

This study focuses on the notion of digital literacy and how it is discussed in relation to Deepfakes. It does so by close-reading a selection of popular videos on YouTube related to the phenomenon in which particular attention is paid to their comment sections through a separate critical discourse analysis. By including the public discourse into the (academic) debate on digital literacy as a solution for fighting misinformation on digital platforms, it becomes possible to assess where the onus should lie according to public: with individuals or with the platforms.

The main question this study is interested in exploring is how YouTube commenters talk about the phenomenon of Deepfakes and whether digital literacy skills are put forward to validate such digitally manipulated content. In order to investigate this topic, this study will first assess what digital literacy entails according to literature and whether the associated skills can indeed be of aid for identifying Deepfakes. After putting digital literacy to the test through a hands-on assessment for three Deepfake videos, this study will consider the comment sections for five videos in total to identify the strategies and solutions that are suggested by users for falsifying Deepfakes. Another subquestion this study explores is how commenters think of such individual strategies for validating content versus regulatory or technical solutions, that put more responsibility on YouTube.

The Deepfakes videos and their comments sections are studied through a textual analysis of popular videos and a critical discourse analysis of their 'top comments'. The method and its limitations will be discussed in chapter 3, alongside the definition of the corpus. Before getting to the analysis, we will first discuss the concept of digital literacy from a theoretical perspective, identifying the associated skills for assessing digital media along with its critiques. This will inform the textual analysis of the Deepfake videos and helps to later contextualize users' comments within the greater discourse on digital literacy. By testing whether the digital literacy skills that are put forward are effective at identifying Deepfakes, we can assess whether the public is doubling down the right skills for validating digital content.

We will also consider current research on YouTube, its business model, and its recommendation system, offering insights into the workings of the platform from an economic and a technical perspective. The followup section will conceptualize YouTube as a discursive space where

[9] Paul Mihailidis and Samantha Viotty. "Spreadable Spectacle in Digital Culture: Civic Expression, Fake News, and the Role of Media Literacies in "Post-Fact" Society." *American Behavioral Scientist*, 61.4 (2018): 441–454.
[10] danah boyd. "You Think You Want Media Literacy… Do You?." *Points* (blog), *Data & Society Research Institute,* March 9, 2018. https://points.datasociety.net/you-think-you-want-media-literacy-do-you-7cad6af18ec2 (accessed March 10, 2019).
[11] Monica Bulger & Patrick Davison.*The Promises, Challenges, and Futures of Media Literacy*. New York: Data & Society Research Institute, 2018, 5. https://datasociety.net/pubs/oh/DataAndSociety_Media_Literacy_2018.pdf (accessed February 13, 2019).

public discourse can take place, along with its limitations. Next to contextualizing the analysis, the theoretical framework will to substantiate the thesis that digital literacy, as an individual skill, is an uncertain answer for dealing with the question of harmful misinformation and manipulated media.

2. Theoretical Framework: Digital Literacy and YouTube

2.1 Defining Digital Literacy

The notion of media literacy has traditionally been described as the critical inquiry into the messages we receive: "the ability to access, analyze, evaluate, create, and act using all forms of communication."[12] Digital literacy can be specified as the "awareness, attitude and ability of individuals to appropriately use digital tools and facilities to identify, access, manage, integrate, evaluate, analyse and synthesize digital resources."[13] In other words, digital literacy is not just about being knowledgeable about digital media, it's about being able to use the features of digital media to examine digital content. Digital literacy includes people's understanding of how knowledge is constructed and expressed through digital media, and the ability to reflect upon these processes.[14] Friesem goes as far as saying that the verification of digital media requires the ability to articulate the political, economic and philosophical motivations of the source.[15] Walker lists a number of techniques for assessing the credibility of manipulated media, which can include Deepfakes: pay attention to the mouth and the eyes to spot unnatural movements, analyse audio separately, assess the motivations for publishing the material.[16] The last follows Friesem, who also stresses the importance of taking the broader context into account.[17] Indeed, it is often the combination of content and context that makes whether a message is accepted as genuine — not just what was shared but also by whom.

Understanding digital literacy as the awareness and ability of individuals to validate digital media is helpful for the analysis of the comment sections beneath the studied Deepfake videos on YouTube. It provides us with cues with what to look for in the comments — basically, all comments that (indirectly) reference individual skills or strategies — to understand whether YouTube commenters (unknowingly) take digital literacy as solution against harmful manipulated content. The general strategies and techniques that are referenced for validating digital media will also be applied to the studied Deepfake videos themselves. This will provide us with insight into whether the commonly suggested digital literacy techniques apply to more advanced forms of manipulated content as well. In a discussion of the analysis, we might then also be able to tell whether the digital

[12] Monica Bulger & Patrick Davison. *The Promises, Challenges, and Futures of Media Literacy*. New York: Data & Society Research Institute, 2018, 7. https://datasociety.net/pubs/oh/DataAndSociety_Media_Literacy_2018.pdf (accessed February 13, 2019).

[13] Allen Martin. "Literacies for the digital age." In *Digital Literacies for Learning*. Ed. Allen Martin and Dan Madigan. London: Facet, 2006, 3–25.

[14] Ibid.

[15] Yonty Friesem. "Teaching Truth, Lies and Accuracy in the Digital Age: Media Literacy as Project-Based Learning." *Journalism & Mass Communication Educator*. 00.0 (2019): 1–14.

[16] Anna Shoenfeld Walker. "Preparing Students for the Fight Against False Information With Visual Verification and Open Source Reporting." *Journalism & Mass Communication Educator*. N.A. (2019): 1–13.

[17] Yonty Friesem. "Teaching Truth, Lies and Accuracy in the Digital Age: Media Literacy as Project-Based Learning." *Journalism & Mass Communication Educator*. 00.0 (2019): 1–14.

literacy strategies that commenters suggest, indeed, can or can not be effective for validating potential Deepfakes.

2.2 Critiquing Digital Literacy as an Individual Solution for the Masses

In "What Hath We Wrought?" boyd questions media literacy as an effective tool in the battle against misinformation. Boyd agrees that, at its core, media literacy is a form of critical thinking that requires people to question the media they receive. It asks for skills at the individual level for validating a piece of information. In today's individualized society, "we double down on media literacy as the 'solution' to misinformation." Boyd questions whether all individuals are able to develop sufficient media literacy skills, noting the different social, economic and political contexts people live in that strongly influence whether people perceive messages as genuine.[18]

While digital literacy specifically refers to knowledge about digital media and also concerns the practical abilities individuals have to validate information, boyd's critique on media literacy advocacy also applies to digital literacy. The critique is indicative for a broader academic debate that is not just about *how* digital literacy can weapon individuals but also about *if* digital literacy should have so much emphasis in society's fight against misinformation. This is not just an academic question as more and more money is invested in training individuals to assess the credibility of the information they consume.[19] Following boyd's sceptical perspective, this study aims to add to the debate by investigating how individuals think about digital literacy. By examining Deepfake videos and whether commenters think such videos should be validated at the individual level, we will be able to tell whether people see digital literacy skills as an effective remedy against misinformation. By bringing the public discourse into the academic discussion, this study aims to offer a more comprehensive overview of the debate.

2.3 Content Circulation in the Age of Algorithms

The circulation of media is increasingly defined by the ways in which information is spread and consumed online.[20, 21] Jenkins, Ford & Green describe a culture that is predominantly consuming media through platforms like Google, Facebook and YouTube as a "participatory culture."[22] In such

---

[18] danah boyd. "You Think You Want Media Literacy… Do You?." *Points* (blog), *Data & Society Research Institute,* March 9, 2018. https://points.datasociety.net/you-think-you-want-media-literacy-do-you-7cad6af18ec2 (accessed March 10, 2019).
[19] Monica Bulger & Patrick Davison.*The Promises, Challenges, and Futures of Media Literacy.* New York: Data & Society Research Institute, 2018, 5. https://datasociety.net/pubs/oh/DataAndSociety_Media_Literacy_2018.pdf (accessed February 13, 2019).
[20] José van Dijck, Thomas Poell & Martijn de Waal. *The Platform Society.* New York: Oxford University Press, 2018.
[21] José van Dijck. *The Culture of Connectivity.* Oxford: Oxford University Press, 2013.
[22] Henry Jenkins, Sam Ford & Joshua Green. *Spreadable Media: Creating Value and Meaning in a Networked Culture.* New York: New York University Press, 2013.

a culture, people can directly engage with the content they consume by, amongst other things, sharing, liking and commenting on the content. Van Dijck, Poell & De Waal go as far as claiming that these platforms have changed societal structures as a whole, defining our current society as the "platform society."[23] In a platform(-based) society, any kind of content — including harmful content — can spread without the help of traditional media through personal channels and platforms.

Today, media circulation and consumption is driven by algorithms that are optimized for maximizing engagement on the advertising-supported platforms that provide many people with news, media and information.[24, 25, 26, 27] Such algorithms allow harmful content to spread more easily, especially when it inflames or aligns with controversial views, which increases engagement and in extent circulation.[28] This became even more apparent after the 2016 U.S. presidential election, during which actors leveraged the technical features of the platforms and their algorithms to strategically help spread misinformation by targeting specific groups with inflaming (clickbait) content on political opponents.[29] The facilitating role platforms play in enabling the circulation of misinformation has increased criticism against the prevailing digital platforms, who now promise to do more to protect users against harmful content.[30, 31, 32, 33]

The critique also concerns YouTube and in particular the working of its monetization and recommendation system. This system is aligned with YouTube's business model which is centred around advertising that supports the content on the platform. Like most of online advertising, advertisers on YouTube are charged based on the number of impressions their ad gets.[34] For users that monetize their content, this means that they earn more money when their content gets more views. This means that both YouTube and video makers are incentivized to get as many people as possible to watch content. YouTube's algorithm plays an important role in this system as 70% of all new views on YouTube are the result of algorithmic recommendations.[35]

[23] José van Dijck, Thomas Poell & Martijn de Waal. *The Platform Society*. New York: Oxford University Press, 2018.
[24] Eli Pariser. *The Filter Bubble: What the Internet is Hiding from You*. New York: Penguin, 2011.
25 Nick Couldry and Joseph Turow "Advertising, Big Data and the clearance of the public realm: Marketers' new approaches to the content subsidy." *International Journal of Communication*. 8 (2014): 1710–1726.
[26] danah boyd. *It's complicated*. New Haven: Yale University Press, 2014.
[27] Thomas Poell and José van Dijck. "Social media and new protest movements." In *The SAGE Handbook of Social Media*. Ed. J. Burgess, A. Marwick, & T. Poell (Los Angeles: SAGE reference, 2018), 546–561.
[28] Cass Sunstein. *Republic.com*. Princeton: Princeton University Press, 2001, 65–67.
[29] José van Dijck, Thomas Poell & Martijn de Waal. *The Platform Society*. New York: Oxford University Press, 2018.
[30] Laura Reed and danah boyd. *Who Controls the Public Sphere in an Era of Algorithms? Questions and Assumptions*. New York: Data & Society Research Institute, 2016. https://datasociety.net/pubs/ap/QuestionsAssumptions_Background-primer_ 2016.pdf (accessed February 10, 2019).
[31] Robyn Caplan and danah boyd. *Who Controls the Public Sphere in an Era of Algorithms*. New York: Data & Society Research Institute, 2016. https://datasociety.net/pubs/ap/MediationAutomationPower_2016.pdf (accessed February 10, 2019).
[32] danah boyd. "Google and Facebook Can't Just Make Fake News Disappear." *Points* (blog), *Data & Society Research Institute,* March 27, 2017. https://points.datasociety.net/google-and-facebook-cant-just-make-fake news disappear-48f4b4e5fbe8 (accessed February 10, 2019).
[33] Hunt Allcott, Matthew Gentzkow and Chuan Yu. "Trends in the Diffusion of Misinformation on Social Media." Working Paper, 2018. https://web.stanford.edu/~gentzkow/research/fake-news-trends.pdf (accessed February 19, 2019).
[34] Karin van Es. "YouTube's Operational Logic: "The View as Pervasive Category." *Television and New Media,* N.A. (2019): 1–17.
[35] Joan E. Solsman. "YouTube's AI is the puppet master over most of what you watch." *CNET,* January 10, 2018 (accessed May 30, 2019).

Several studies have revealed that YouTube's algorithm recommends 'extreme content' disproportionately to how popular these videos look to be in terms of quantitative engagement measures (views, comments, shares, etc.).[36, 37] Rieder et al. found that YouTube search results prefer controversial, YouTube-native content over mainstream media content, even though these videos have received less views.[38] Tufekci has called YouTube "the great radicalizer," arguing that YouTube has learned that users are more likely to stay on the platform when they are recommended extreme, radicalizing content. Tufekci flags the responsibility YouTube has towards users, arguing that society should not accept a company that makes so much profit while potentially radicalizing people because it gets them to look at more and more extreme content through its recommendation system.[39]

While the shift in media circulation away from traditional gatekeepers is not inherently harmful, theory as well as empirical research into the workings of platforms and their business models suggest that harmful content can spread more easily through recommendation systems. Many people will not be fully aware of how the content of their personalized recommendations, feeds and search results are constituted, and what incentives are at play. According to Gillespie, the algorithms that account for an increasingly large portion of what we see are still something "unattainably complex" to the public.[40] Digital literacy can help to make people more aware of today's algorithmic-driven media circulation, and the kind of content that is more likely to be recommended on platforms like YouTube. Digital literacy could, for example, make people aware that recommended content is not necessarily editorially endorsed or popular on the platform in a traditional sense. It will be instructive to see to what extent users show to be aware of YouTube's workings, and whether commenters (indirectly) refer to digital literacy skills as a way of avoiding the harmful kinds of content YouTube favours in its recommendations.

2.4 YouTube as a Discursive Space

Recommendation systems have not only made it easier for content to spread. Pariser claims that personalized recommendation systems cause the public to consume media in an increasingly fragmented way. According to Pariser, this had led to a fragmentation of the traditional public sphere.[41] This is not to say there is no public sphere on YouTube. According to Burgess & Green,

---

[36] Bernhard Rieder, Ariadna Matamoros-Fernández, and Òscar Coromina. "From Ranking Algorithms to 'Ranking Cultures': Investigating the Modulation of Visibility in YouTube Search Results." *Convergence*, 24.1 (2018): 50–68.

[37] Zeynep Tufekci. "YouTube, the Great Radicalizer." *The New York Times,* March 10, 2018, https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html (accessed May 3, 2019).

[38] Bernhard Rieder, Ariadna Matamoros-Fernández & Òscar Coromina. "From Ranking Algorithms to 'Ranking Cultures': Investigating the Modulation of Visibility in YouTube Search Results." *Convergence*, 24.1 (2018): 50–68.

[39] Zeynep Tufekci. "YouTube, the Great Radicalizer." *The New York Times,* March 10, 2018, https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html (accessed May 3, 2019).

[40] Tarleton Gillespie. "Algorithm." In *Digital Keywords: A Vocabulary of Information Society and Culture*. Ed. by Benjamin Peters. Princeton, NJ: Princeton University Press, 2016, 28–31.

[41] Eli Pariser. *The Filter Bubble*. The Penguin Press: New York, 2011.

YouTube is a "cultural public sphere [...] a communicative space and a community" that exemplifies participatory media with its potential for civic exchange.[42]

The notion of a public sphere refers to Habermas, who used it to describe "a discursive space in which members of a community can discuss important matters of the day."[43] YouTube can be seen as cultural public sphere because it's "an enabler of encounters with cultural differences and the development of political 'listening' across belief systems and identities."[44] Indeed, because the content plays a key role in determining the audience for a video, YouTube videos and their comments can constitute a public sphere that differs from the traditional notion of a (predefined) community. The Deepfake videos and comments that are the subject of this study will allow us to get insights into the broader public discourse on the phenomenon of Deepfakes. The comments can be seen as an indirect exchange of views that provides insight into what certain members of the public claim to think about the phenomenon and how its spread can or can not be mitigated.

While the public sphere has been fragmented, YouTube videos and their comments allow for a new way to get insight into the public discourse. While this kind of analysis comes with its own limitations, it is arguably also richer than the traditional public sphere of Habermas, which falls short at capturing the heterogeneity of society. According to Reed & boyd, digital platforms like YouTube are rarely homogenous on issues of public concern — such as the phenomenon of Deepfakes.[45]

YouTube's community differs from that of other platforms as its content-based audiences are generally less cohesive and active than those of person-centred platforms like Facebook, Instagram and Tumblr.[46] While users can directly interact with each other through social-networking features such as threaded comments and comment likes, Rotman, Golbeck & Preece argue this does not necessarily provide YouTube with a strong sense of community. Their study on YouTube's comment culture confirms there are relatively few interactions between users in the comment sections.[47] The apparent lack of direct interactions between users suggests that most comments should be seen as a user's reaction to the video, and much less so as a response to other members of the community. There is no singular public debate beneath videos, and relatively few people add to the debate. One of the implications of this is that you can not expect to find consensus in the traditional sense of the word — though one might very well find a majority of YouTube commenters individually agree on a topic. Another limitation that needs to be taken into account is that YouTube

[42] Jean Burgess & Joshua Green. *YouTube: Online Video and Participatory Culture*. Malden: Polity Press, 2009, 69.
[43] Laura Reed & danah boyd. Who Controls the Public Sphere in an Era of Algorithms? Questions and Assumptions. New York: Data & Society Research Institute, 2016. https://datasociety.net/pubs/ap/QuestionsAssumptions _Background-primer_2016.pdf (accessed February 10, 2019).
[44] Jean Burgess & Joshua Green. *YouTube: Online Video and Participatory Culture*. Malden: Polity Press, 2009, 77.
[45] Laura Reed & danah boyd. Who Controls the Public Sphere in an Era of Algorithms? Questions and Assumptions. New York: Data & Society Research Institute, 2016. https://datasociety.net/pubs/ap/QuestionsAssumptions _Background-primer_2016.pdf (accessed February 10, 2019).
[46] Dhiraj Murthy & Sanjay Sharma. "Visualizing YouTube's comment space: online hostility as a networked phenomena." *New Media & Society*, 21.1 (2019): 191–213.
[47] Dana Rotman, Jennifer Golbeck & Jennifer Preece. "The community is where the rapport is – on sense and structure in the YouTube community." In: *Proceedings of the fourth international conference on Communities and Technologies (C&T '09), University Park, PA, 25–27 June 2009*. New York, NY: ACM, 2009, 41–50.

commenters only present a fraction (often less than one percent) of the viewers of a video. One needs to extra careful with generalizing findings from the comments for the audience group as a whole. This is in part because YouTube comments are typically not representative as they do not represent a fair cross-section of society or of the complete audience group for a video. It takes a kind of person and video subject, to participate to a video by leaving a response, let alone start a discussion with other YouTube members they might not even know.[48]

[48] Mike Thelwall, Pardeep Sud & Farida Vis. "Commenting on YouTube videos: From guatemalan rock to El Big Bang." *Journal of the Association for Information Science and Technology.* 63.3 (2012): 616-629.

3. Method: Close-reading Deepfake Videos and Their Comments

To better understand the public discourse on Deepfakes through YouTube, and whether commenters refer to digital literacy skills as a tool to mitigate the harmful potential of such videos, this study combines a textual analysis of Deepfakes videos on YouTube with a critical discourse analysis of their comments. The textual analysis, or close reading, of three popular examples of Deepfake videos will provide us with a more concrete and detailed understanding of the subject matter. This will later allow us to assess whether the digital literacy skills that are mentioned in response to Deepfakes are effective at validating such content, or whether the public, too, is doubling down on ineffective techniques for fighting misinformation, as the critical discourse analysis of the comment sections will point out.[49]

3.1. Corpus Selection

The corpus for this study consists of five popular Deepfake videos on YouTube and their top (500) comments as defined by the platform. Two of the videos feature discussions of the technology and have been selected as part of the corpus because they're amongst the most watched and commented on videos on Deepfakes on YouTube. The other videos are examples Deepfakes that feature manipulated statements from high-profile figures such as former U.S. President Barack Obama, Russian President Vladimir Putin, and Facebook CEO Mark Zuckerberg. Next to a critical discourse analysis of their top comments, these three videos will also be examined through a textual analysis. The corpus is limited to five videos and their highest-ordered comments with a maximum of 500 responses per video to not completely skew the results of the analysis to the findings from the most popular videos. Because some videos have considerably fewer comments than other videos, the corpus for this part of the study consists of 1,742 comments (Table 1).

Video Title	Uploader	Views	Like/Dislikes	Comments Total/Analysed
You Won't Believe What Obama Says In This Video! 😉	BuzzFeedVideo	5,988,018	78,722 / 13,638	8,393 / 500
Face2Face: Real-time Face Capture and Reenactment of RGB Videos (CVPR 2016 Oral)	Mattias Niessner	4,584,887	25,422 / 667	2,303 / 500
It's Getting Harder to Spot a Deep Fake Video	Bloomberg	2,323,564	41,584 / 2,018	2,931 / 500

[49] danah boyd. "You Think You Want Media Literacy… Do You?." *Points* (blog), *Data & Society Research Institute,* March 9, 2018. https://points.datasociety.net/you-think-you-want-media-literacy-do-you-7cad6af18ec2 (accessed March 10, 2019).

Putin | Face Replacement	derpfakes	69,299	329 / 16	65 / 65
A WORLD WITHOUT FACEBOOK	Bayview Drive Films	62,112	376 / 164	177 / 177

Table 1. Video details for analysed material as of 22 April, 2019.

The comments analysed in this study were retrieved using the "Video Info and Comments Module" of the Digital Methods Initiative.[50] This tool provides spreadsheets with the full metadata of YouTube videos including their entire comments section including subthreads, which can manually only be retrieved by scrolling to load additional comments in small batches. Because the tool is unable to indicate how comments are ranked by YouTube's 'top comments' ordering, the (500) top comments for each video still had to be filtered manually by cross-checking the overview with the usernames behind the top (500) comments on YouTube.

Because for most YouTube videos, the majority of comments are written in English, comments in other languages are excluded from the analysis, which would otherwise have resulted in language and interpretation problems.[51] In total, only 82 of the 1,742 top comments were not written in English and excluded from further analysis. The final dataset for the critical discourse analysis comprised 1,660 top comments along with 1,418 additional, threaded replies, receiving 68,460 likes in total. The way these 3,078 total comments are distributed across the five videos presents a limitation of the study, as the comments are not evenly spread between the videos. For two videos, all comments are considered. For the three other videos, less than a quarter of all comments are included in the analysis yet these videos account for 86% of all comments in the corpus. In order for the comment corpus to have sufficient substance, it is skewed towards the most popular videos (albeit not completely). While it is still possible to make generalized claims about the comments of YouTubers on Deepfake videos, the uneven distribution of the corpus has to be minded.

3.2. Textual Analysis of Deepfake Videos

The textual analysis of three Deepfake videos means close reading each video to focus on the artificial or manipulated elements, most notably the head, mouth and eyes of the featured leaders: Obama, Putin, and Zuckerberg. These 'moving parts' are most likely to be manipulated, especially through Deepfakes, and paying close attention to these elements is one most frequently mentioned digital

---

[50] Bernard Rieder. "Video Info and Comments Module." *YouTube Data Tools*. N.D.
https://tools.digitalmethods.net/netvizz/youtube/mod_video_info.php (accessed April 22, 2019).
[51] Dana Rotman, Jennifer Golbeck & Jennifer Preece. "The community is where the rapport is – on sense and structure in the YouTube community." In: *Proceedings of the fourth international conference on Communities and Technologies (C&T '09), University Park, PA, 25–27 June 2009*. New York, NY: ACM, 2009, 41–50.

literacy strategies for spotting manipulated content. Another content verification technique mentioned by Walker that is applied is analysing the audio of the videos separately.[52] In addition to being able to identify particularities more easily, this part of the analysis also allows us to pay particular attention to the message of each video and identify its purpose.

The textual analysis of the videos offers insight into the current level of sophistication Deepfake videos offers, and whether the concerns about the potential of this technology are justified. The analysis serves as a background for the critical discourse analysis of the YouTube comments. It not only points out the differentiating features of content that is created with the Deepfake technology, it also offers a way to test the recommended digital literacy techniques themselves. In other words, it allows us to validate whether the validation strategies that are put forward are effective at identifying Deepfakes.

3.3. Critical Discourse Analysis of YouTube Comments Related to Deepfakes

The critical discourse analysis of YouTube comments offers a way to systematically analyse discursive forms to uncover patterns and relationships.[53] The analysis involves a close-reading of the selected comments, in which particular attention is paid to the political aspects of the comments — taking into account the broader context in which things are said. In the analysis of comments, the number of likes comments have amassed is also taking into account as a natively digital way for YouTube users to support the statements of other users — 'adding' weight to the discussion without articulation. By paying particular attention to the comments that other users have given extra importance, the analysis allows us to identify how YouTube users think about Deepfakes and whether individuals can or can not be trained to detect them.

The critical discourse analysis comprises three levels. The first level is general reading of the comments as a regular viewer on YouTube.[54] This reveals the general sentiment and types of discussions users engage in relation to Deepfakes. The preliminary analysis of the comments informs the categorizing and coding system for a more for the more structural analysis at the second level.

The second level of analysis pays particular attention to whether the topics that are put forward in academic literature on digital literacy and the spread of misinformation are also mentioned by the public on YouTube. In other words, do users talk about video-assessment techniques, do they reflect upon how knowledge is constructed through digital media, and do they discuss YouTube's

[52] Anna Shoenfeld Walker. "Preparing Students for the Fight Against False Information With Visual Verification and Open Source Reporting." *Journalism & Mass Communication Educator*. N.A. (2019): 1–13.
[53] Thomas Huckin. "Critical discourse analysis and the discourse of condescension." In E. Barton and G. Stygall (Eds.), *Discourse Studies in Composition*. Cresskill, NJ: Hampton Press, 2002, 155–176.
[54] Daren C. Brabham. "THE MYTH OF AMATEUR CROWDS." *Information, Communication and Society*, 15.3 (2012): 394–410.

role in the spread of manipulated content? Because the majority of comments are only a couple words or a sentence long, each comment is categorized under one of 12 (sub)categories:

Deepfakes 1.1	Concerns about Deepfakes, for example users pointing out that the technology can be dangerous.
Deepfakes 1.2	Enthusing about Deepfakes, for examples commenters that indicate they want to start making Deepfakes themselves
Deepfakes 1.3	Ridiculing Deepfakes: comments that say the public should not worry about the technology, for example by making jokes about it.
Digital literacy 2.1	Statements that the video in question is fake, without giving further explanation.
Digital literacy 2.2	References to the facial expressions of the person in the video, particularly focusing on the person's eye and lip movements
Digital literacy 2.3	References to the sound of a video as a way to identify the video as fake.
Digital literacy 2.4	The commenters in this category both mention the facial expressions of the person in the video and the sound.
Digital literacy 2.5	A category of commenters who express their distrust in a video's source.
Regulatory 3.1	Commenters who say that YouTube has to do something against Deepfakes, for example using technical means to prevent the spread of these videos.
Regulatory 3.2	Comments by users who say that governments have to prohibit the technology in order to stop these videos from being published.
Regulatory 3.3	Comments by users mentioning technical methods to verify Deepfakes.
Miscellaneous 4	Comments that did not fit into any of the categories above.

Table 2. Comment categorisation (sub)categories.

The last level of analysis constitutes a reflection on the discourse based on the author's own understanding of the discursive objects.[55] Here, we can connect what is said by the public on YouTube to what is said by scholars in academic literature to analyse whether their views on digital literacy are aligned in relation to the phenomenon of Deepfakes.

---

[55] Daren C. Brabham. "THE MYTH OF AMATEUR CROWDS." Information, Communication and Society, 15.3 (2012): 394–410.

4. Findings: Deepfakes and Their Discourse on YouTube

In this section, we will first present the findings from the textual analysis of three Deepfake videos. This is followed by the critical discourse analysis of YouTube comments through the three levels of analysis as presented in the method section.

4.1 Textual Analysis: Debunking Deepfakes through Digital Literacy

The analysis of three Deepfake videos found that verification techniques that are put forward by digital literacy are indeed effective for identifying them as fabrications. The findings are discussed separately per video in the next sections.

4.1.1 "President Trump is a Total and Complete Dipshit"

"We're entering an era in which our enemies can make anyone say anything at any point in time," says Barack Obama in a popular Deepfake demonstration video from April 2018. In the video, Obama warns the public about the new manipulation technology, saying people should not believe everything they see and hear online, closing with the note "stay woke bitches."[56] In no small part because of the lifelike voice-dubbing by comedian Jordan Peele, the combination of image and sound comes across as believable at first. It is the contents of Obama's message that indicates that the video is not original. Indeed, upon closer look, it is noticeable that the bottom edges of Obama's face are not as sharp as the other edges in the video. The contours of his lower face are blurry and slightly asymmetrical, while the folds around Obama's mouth are sharp but do not follow through beneath the mouth (fig. 1).

---

[56] BuzzFeedVideo. "You Won't Believe What Obama Says In This Video! 😉." YouTube video, 1:12, April 17, 2018. https://www.youtube.com/watch?v=cQ54GDm1eL0 (accessed April 22, 2019).

Figure 1. Comparison of the Obama Deepfake on the left and the original shot on the right.[57, 58]

The audio of the video also does not hold up upon closer inspection. On its own, Obama's voice sounds reasonably genuine but when compared to the original fragment, the discrepancies are also noticeably for the untrained ear, with Peele's impersonation sounding significantly more nasal than Obama himself. The validation techniques put forward by Walker — looking for unnatural facial features and analysing the audio separately — here are effective for confirming the BuzzFeed video is manipulated.[59]

Looking beyond the technical features of the video, the intent of the video and purpose of the source are quite clear in this case.[60] The BuzzFeed Deepfake takes meta-approach by making Obama say that certain people can now make him say anything. The purpose of the video is not to fool the viewer but to educate them. This is emphasized when the video continues in split screen, showing Peele dubbing the Deepfake Obama in perfect sync. It is not necessary to confirm through other sources that the statements in the video are not real because this is already revealed in the video itself. However, if viewers were to Google the statement that "President Trump is a total and complete dipshit," they would find that all of the top search results indicate that the statement is fake.

4.1.2 Putin Faceswapped

The second Deepfake video analysed in this study stars Vladimir Putin alongside a clip for *Saturday Night Live* on which the Deepfake is based.[61] The video was uploaded by derpfakes, a popular

---

[57] BuzzFeedVideo. "You Won't Believe What Obama Says In This Video! 😉." YouTube video, 1:12, April 17, 2018. https://www.youtube.com/watch?v=cQ54GDm1eL0 (accessed April 22, 2019).

[58] The Obama White House. "President Obama Speaks to the Open Government Partnership Global Summit." YouTube video, 2:09, December 17, 2016. https://www.youtube.com/watch?v=-mIaZPqoOYw (accessed June 13, 2019).

[59] Anna Shoenfeld Walker. "Preparing Students for the Fight Against False Information With Visual Verification and Open Source Reporting." *Journalism & Mass Communication Educator.* N.A. (2019): 1–13.

[60] Yonty Friesem. "Teaching Truth, Lies and Accuracy in the Digital Age: Media Literacy as Project-Based Learning." *Journalism & Mass Communication Educator.* 00.0 (2019): 1–14.

[61] derpfakes. 2018. "Putin | Face Replacement." YouTube video, 0:37, Feb 26, 2018. https://www.youtube.com/watch?v=hKxFqxCaQcM&feature=youtu.be (accessed April 22, 2019).

YouTube channel dedicated to Deepfakes that is famous for making some of the Nicholas Cage Deepfake memes.[62, 63] Like most of the channel's videos, the Putin Deepfake is posted for demonstrational purposes rather than for manipulative purposes.[64] The face of *SNL*-actor Beck Bennett is swapped for Putin's but the audio is still Bennett's original. The split-screen setup and bare-chest setting of the video makes no attempt at suggesting the clip of Putin is genuine.

Because both clips are juxtaposed, it is easy to note the discrepancies between the original content and the manipulated footage. When you look at Putin's face, you can notice that Putin's entire face is added onto the original footage, instead of that facial features such as the mouth are digitally manipulated, as was the case for the BuzzFeed video. Especially around the eyebrows, you can see brief glitches in the composition of the added-on face. You can also spot the Deepfake by looking at the mouth, where Putin's teeth naturally appear and disappear when opening his mouth (fig. 2).[65] While the juxtaposition arguably makes it unnecessary to confirm that the footage of Putin is manipulated, digital literacy techniques can quickly corroborate his.

Figure 2. Screenshot of the Putin video with *SNL*-actor Beck Bennett on the left and on the right a version where Bennett's face is swapped for Putin's.[66]

[62] derpfakes is also one of the moderators of the r/GiFFakes subreddit in which users share the(ir) latest Deepfake creations.

[63] Sam Haysom. "People are using face-swapping tech to add Nicolas Cage to random movies and what is 2018." *Mashable,* January 21, 2018. https://mashable.com/2018/01/31/nicolas-cage-face-swapping-deepfakes/?europe=true (accessed June 3, 2019).

[64] derpfakes. "10k Subscriber Deepfakes Special | Derpfakes." YouTube video, 1:59, February 8, 2019. https://www.youtube.com/watch?v=UbS8XKzwECA (accessed May 26, 2019).

[65] Konrad Weber. "8 steps to verify Deep Fake videos." *Medium* (blog), *Konrad Weber,* November 8, 2018. https://medium.com/@konradweber/8-steps-to-verify-deep-fake-videos-1dfc408568c4 (accessed May 25, 2019).

[66] derpfakes. 2018. "Putin | Face Replacement." YouTube video, 0:37, Feb 26, 2018. https://www.youtube.com/watch?v=hKxFqxCaQcM&feature=youtu.be (accessed April 22, 2019).

4.1.3 A World Without Facebook

Mark Zuckerberg announces that he will delete Facebook for good because of the concerns about Facebook's negative impact on people's lives.[67] This is what viewers see in the video by filmmaker Andrew Oleck that was uploaded on April Fools' Day 2018. The video was posted on YouTube as well as on Facebook where it received 32 million views in 18 days, and over 70 million views as of June 2018.[68] Intended as a joke, the announcement was formally debunked by fact-checkers of *Snopes*, who published an article saying that many online viewers thought Zuckerberg's statements were authentic.[69]

Despite that the message and language in the video are extremely over the top for an official CEO announcement — especially when compared to how Zuckerberg talks about himself and Facebook in reality — it is also clear from the technical features of the video that the content has been digitally manipulated. Like the other two videos analysed by this study, the manipulated content is based on a static shot of the public figure in a sitting position. The most important part of the footage that needs to be manipulated in order to have Zuckerberg believably say something different than the original is the mouth. And indeed, it is easily possible to see that the mouth has been added onto the original image, the rest of which looks untouched compared to the original footage (fig. 3).[70] This suggests that Zuckerberg video may actually not be a Deepfake generated by software, but a traditionally manipulated video in which the mouth is swapped for that of a (voice) actor. This is further emphasized by the facial expressions and movements of Zuckerberg's throat, which are not in sync with his mouth movements. Lastly, by comparing the audio to an original fragment of Zuckerberg's voice, you can hear that Zuckerberg's voice in the video is noticeably lower — though by itself, the impression reasonably convincing.

[67] Bayview Drive Films. "A WORLD WITHOUT FACEBOOK." YouTube video, 3:23, April 1, 2018. https://www.youtube.com/watch?v=ADCPqE4ShGQ#action=share (accessed April 22, 2019).
[68] Video of The Internet. "A WORLD WITHOUT FACEBOOK." Facebook video, 3:23, Apr 23, 2018. https://www.facebook.com/VIDEOOTI/videos/352738545218184/ (accessed April 22, 2019).
[69] Dan Evon. "Did Mark Zuckerberg Announce That He's Closing Facebook?" *Snopes*, April 18, 2018. https://www.snopes.com/fact-check/mark-zuckerberg-closing-facebook/ (accessed June 3, 2019).
[70] PBS NewsHour. "WATCH: Facebook's Mark Zuckerberg to cooperate on Russian investigation after 2016 election." YouTube video, 8:31, September 21, 2017. https://www.youtube.com/watch?v=wsUhyiO31DY (accessed June 3, 2019).
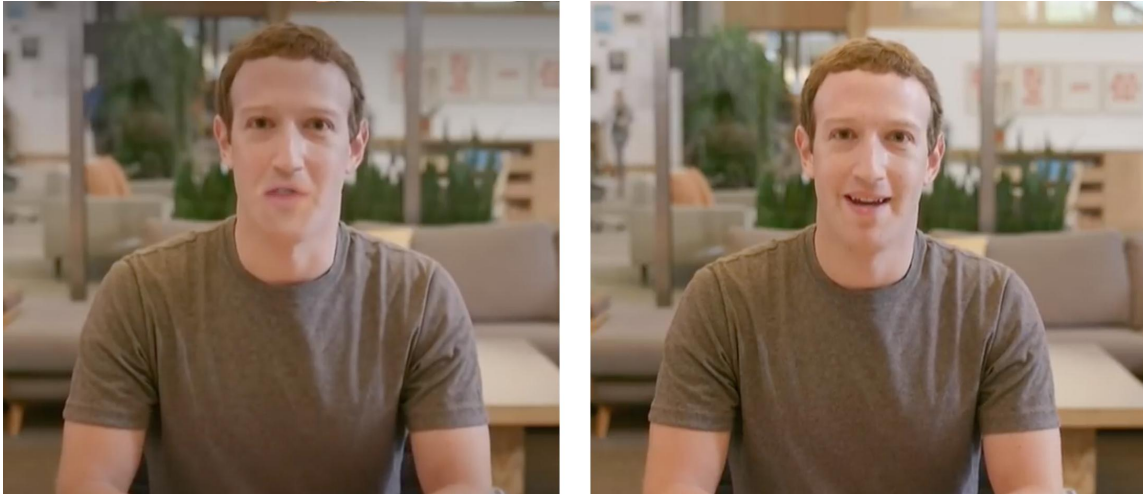
Figure 3. Comparison of the manipulated Zuckerberg on the left and the
original footage on the right.[71, 72]

4.2 Critical Discourse Analysis: Commenting on Deepfakes and Digital Literacy

The first level of analysis found that 38% of the comments did not concern any of the 11 main topic categories outlined in the methods section (table 2). As a result, 1,180 out of 3,077 comments are classified as miscellaneous and not considered for further analysis. The categorisation of the remaining 1,897 comments that are under analysis (at the second level of the critical discourse analysis) show that 52% of the comments on the (5) Deepfake videos are focussed on the phenomenon itself. 35% of the comments can be seen to refer to digital literacy or skills, while many of such comments are in response to the technology of Deepfakes. The remaining 13% of the comments relate to possible regulation in the broadest sense of the word and include solutions against the spread Deepfakes that are not at the individual level of digital literacy.

The findings from the second level of the critical discourse analysis will be presented in the next sections. The discussion of the findings happens at the third level of analysis (chapter 5), in which we will critically reflect on the findings from the discourse analysis.

4.2.1 Deepfakes: A Tool for Manipulation or for Memes?

Of the 978 comments that are about Deepfakes, 56% can be seen as concerned about (the potential of) the technology. However, there is not a lot of agreement between commenters on how the Deepfakes might be misused. A third of the concerns about Deepfakes are alarming. Many of such

---

[71] Bayview Drive Films. "A WORLD WITHOUT FACEBOOK." YouTube video, 3:23, April 1, 2018.
https://www.youtube.com/watch?v=ADCPqE4ShGQ#action=share (accessed April 22, 2019).
[72] PBS NewsHour. "WATCH: Facebook's Mark Zuckerberg to cooperate on Russian investigation after 2016 election."
YouTube video, 8:31, September 21, 2017. https://www.youtube.com/watch?v=wsUhyiO31DY (accessed June 3, 2019).

comments paint a dystopian view of the technology that might even cause World War Three according to 43 individual commenters. 9% of the concerns are about governments, who might use Deepfakes to influence people, while a similar amount of concerns claim that mainstream media will use the technology to spread misinformation amongst the public. This last group of comments shows a notable distrust of mainstream media, most explicitly CNN.

On the other side of the discourse on Deepfakes, 12% of the comments express an interest in the technology without a concern. Such comments often hail Deepfakes' potential for memes and film making without touching upon the harmful potential of such content. Neither do these comments engage or counter the concerns shared by other users. 11% of all comments relating to the topic of Deepfakes even ridicule the concerns about the technology by joking about the concerns. While the first two groups mentioned here — the concerned and unconcerned — can be seen as the traditional two sides of the debate, there is hardly any interaction between these two groups. This is quite different for last group that actively play to (and with) the concerns of other users, which successfully provokes reactions that lead to short discussions, but never to consensus.

4.2.2 Strategies and Solutions: Validating Content

Of the comments that can be seen to reference digital literacy or its related skills, 65% point to audio-visual aspects that reveal the video has been manipulated, most often the mouth. This suggests that the basic content validation techniques proposed by digital literacy are known by a considerable share of users, knowing that fewer than a percent of all viewers leave a comment.[73] Based on the analysis of the videos in question, this study can confirm that the pointers that are shared on YouTube are indeed sufficient to identify the current generation of Deepfakes.

The majority of such comments do not mention digital literacy explicitly but point to a specific aspect of the video and include a (hyperlinked) timestamp so that other users can easily validate the information.[74, 75, 76] Instructive comments that offer a more detailed analysis of the content validation are found to receive a significant amount of support from other users, receiving 20 up to 350 likes. Most YouTube comments do not receive any likes, while the short comments that

---

[73] Mike Thelwall, Pardeep Sud & Farida Vis. "Commenting on YouTube videos: From guatemalan rock to El Big Bang." *Journal of the Association for Information Science and Technology.* 63.3 (2012): 616–629.

[74] Leadsticks. "0:51 This is all fake. The eyes don't blink at the same time. Still don't believe me? Look at the eyebrows." January 28, 2019. Comment on Bloomberg, "It's Getting Harder to Spot a Deep Fake Video." https://www.youtube.com/watch?v=gLoI9hAX9dw&lc=UgxeXgci5R9wc3KhP-R4AaABAg (accessed April 22, 2019).

[75] William Greene. "It look like Obama got beat up... a swollen left jaw and perhaps maybe a black eye with some make-up covering it up...damn!." December 14, 2018. Comment on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=Ugzvegk2gkebr6ydGQx4AaABAg (accessed April 22, 2019.

[76] *ïï Clorox Bleach ïï*. "Guys this is a joke and edited, I did research and check on his lips.. As you can see it is not his lips but someone else's edited there." April 13, 2018. Comment on Bayview Drive Films, "A WORLD WITHOUT FACEBOOK." https://www.youtube.com/watch?v=ADCPqE4ShGQ&lc=UgzXlUxKktlrnQI1-jF4AaABAg (accessed April 22, 2019).

only name a specific aspect or timestamp receive a maximum of 10 likes, or none. Only one the comments that referred to a digital literacy technique received a written appreciation.[77]

Some commenters can be seen to show off their digital literacy skills by pointing out how fast they saw the video was a Deepfake. These commenters and those that go a step further by suggesting that people are stupid if they can not tell a video is fabricated receive amongst the highest numbers of likes from all comments in this study, receiving over a 1,000 or even over 2,000 likes.[78, 79] What is interesting is that the thread replies to such comments include a fair amount of critique from users that question whether Deepfakes are indeed so stupidly simple to spot for the average person.[80, 81] These threaded critiques hardly ever receive likes from other users, but this could be explained by design. By default, YouTube hides threaded replies behind a button that means that most of the true discussions are out of the view to the majority of users.

Other types of critiques suggest we have not seen the true potential of the technology, claiming that Deepfakes in the future will be far more convincing.[82] In addition to general advances of the technology behind Deepfakes, commenters reference developments in the field of audio manipulation.[83] Such comments imply that some of the digital literacy techniques people currently rely on might not hold up in the future.

35% of the comments that relate to the notion of digital literacy focus on the video's source instead of the video's content. This is especially because of the many critical comments on the videos posted by Bloomberg and BuzzFeed. For example, multiple commenters specifically refer to the irony of BuzzFeed's Deepfake demonstration video, which urges viewers to rely on trusted news sources.

---

[77] Michael Isaac. "Noticed his face especially around mouth is swollen and the voice isn't quite right. Great observation!" April 18, 2018. Comment reply on massessum1, "You can see, if you look closely, that Obama's face is animated, especially his smile lines. Plus, the voice is not quite right. Pretty damn good, though." YouTube comment, April 18, 2018. https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=Ugxnp-hbnr7yMsO9Yap4AaABAg.8fBZ3O89cHG8f BgDFtQIbE (accessed April 22, 2019).

[78] Mauri QHD. "if you are not 100% sure is not him as soon as 0:02, you are part of the problem." April 18, 2018. Comment on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgwdX1dvbY7lZcCrEL14AaABAg.8f8p_CQcL5g8fAAsDu TyOi (accessed April 22, 2019).

[79] G.L. Harrison. "I knew it wasn't him talking straight off the bat." Comment on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgwHsaB v1QuCLBjgYpB4AaABAg (accessed April 22, 2019).

[80] Kairas_love. "Eve[r]yone on here says it looks fake because you know that from the start. The issue being adressed here is that the mass public wouldn't be able to tell. Say this was posted on television and your parent[s] or Grandparents saw this. […]" April 17, 2018. Comment on BuzzFeedVideo. "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgxaLrzhpip9TjJo0ut4AaABAg (accessed April 22, 2019).

[81] Colleen Casey. "Perhaps instead of 'insulting' people who do not dedicate their lives to 'technology' and non-living things, and choose to live in the 'real world of Human Beings'... […]" Comment reply on ThePromisedWLAN. "This is frightening technology for media manipulation," October 7, 2016. https://www.youtube.com/watch?v=ohmajJTcpN k&lc=UgjJzYe0lhU_JHgCoAEC (accessed April 22, 2019).

[82] largelysubatomic. "this isn't even a good example of how effective this technology is right now let alone in the future. this is just a meme vid from buzzfeed while nation states, […]" April 18, 2018. Comment reply on Mauri QHD, "if you are not 100% sure is not him as soon as 0:02, you are part of the problem." April 18, 2018. https://www.youtube.com/ watch?v=cQ54GDm1eL0&lc=UgwdX1dvbY7lZcCrEL14AaABAg.8f8p_CQcL5g8fAAsDuTyOi (accessed April 22, 2019).

[83] Fanny Schmeller. "just wait until adobe release their photoshop for audio software that theyv've been demoing. we wont have a clue whats real." November 27, 2016. Comment reply on ThePromisedWLAN, "This is frightening technology for media manipulation," October 7, 2016. https://www.youtube.com/watch?v=ohmajJTcpNk&lc= UgjJzYe0lhU_JHgCoAEC (accessed April 22, 2019).

Such comments reference the statements in the video, saying things like "Because buzzfeed is a trusted news source loooool," and "Lmao rely on trusted news sources, is not buzzfeed the most distrusted 😂😂😂."[84, 85] Similarly, several commenters make ironic statements in relation to Bloomberg's video that informs viewers about the increased level of sophistication in Deepfakes: "Too funny! Bloomberg will be the first ones to use this new technology to spread their fake news."[86]

Such comments show that at least a part of the YouTube public is critical towards a video's source, which is a key principle of digital literacy.[87] However, it is all but clear whether these commenters genuinely question the video's source, because of digital literacy, or because they do not trust mainstream media in general.[88] The comments are aimed at the messenger instead of the message, and none share research or evidence that substantiates the claim that BuzzFeed or Bloomberg cannot be trusted. A number of such comments draw a somewhat fatalistic response for other users. These commenters share the view that people *should* be critical of video's source but are concerned that many people will simply believe and share videos as long as it aligns with their existing views, echoing Sunstein.[89] One user goes as far as saying that even just *wanting* to believe something in a Deepfake video can lead to fatal consequences.[90]

So, while commenters generally seem to pay attention to the video's source, it is questionable whether the technique is applied in the considered way digital literacy assumes. At the same time, many commenters themselves question whether other people would even consider the video's source when the contents of the message is something they already (want to) believe.

---

[84] Owen Hutchins. "Because buzzfeed is a trusted news source loooool." April 19, 2018. Comment on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc= UgyBF6R1UuTdG1OpUAZ4AaABAg (accessed April 22, 2019).

[85] Aaron Diaz. "Lmao rely on trusted news sources, is not buzzfeed the most distrusted 😂😂😂." April 19, 2018. Comment on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgwSOpLvRgsXRZpIRJh4AaABAg (accessed April 22, 2019).

[86] Meme mine. "Too funny! Bloomberg will be the first ones to use this new technology to spread their fake news." October 17, 2018. Comment on Bloomberg, "Bloomberg. "It's Getting Harder to Spot a Deep Fake Video." https://www.youtube.com/watch?v=gLoI9hAX9dw&lc=UgxQC3UX77HDsm6Ut3B4AaABAg (accessed April 22, 2019).

[87] Example: PUMPA DOUR. "I have never ever believed the mass media before, so this shit is not too scary to me. I will continue not believing anything i see and hear." April 10, 2016. Comment on Matthias Niesner, "Face2Face: Real-time Face Capture and Reenactment of RGB Videos (CVPR 2016 Oral)." https://www.youtube.com/watch?v= ohmajJTcpNk&lc=Ugg2AmnPj8qDcHgCoAEC (accessed April 22, 2019).

[88] danah boyd. "You Think You Want Media Literacy… Do You?." *Points* (blog), *Data & Society Research Institute,* March 9, 2018. https://points.datasociety.net/you-think-you-want-media-literacy-do-you-7cad6af18ec2 (accessed March 10, 2019).

[89] Cass Sunstein. *Republic.com*. Princeton: Princeton University Press, 2001, 65–67.

[90] Ibieiniid. "all it would take to convince some people is for them to WANT to believe what theyre seeing. If there was a grainy video of Obama saying 'I murdered Seth Rich after a pedo party at Comet Ping Pong', whats to say some gullible idiot wouldnt be planning Pizzagate part 2 immediately? Truth is, there ARE some people you can't help, because you're an internet nobody (not an insult, just fact). Jordan Peele has the stage to try to help those people." April 20, 2018. Comment reply on moiraine_damodred, "if you couldn't immediately tell that was computer generated I can't help you." April 19, 2018. https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgyO2NJmqvH4JL861OF4AaABAg.8f Dkcj8tmPb8fGLmUCU_vt (accessed April 22, 2019).

4.2.3 People versus Platforms: Where the Onus Should Lie According to Commenters

Only 7% of comments on regulation refer to the role of YouTube in relation to the spread of Deepfake videos, and the few users that do talk about the platform hardly ever mention the recommendation system in this context. While studies have highlighted the role of YouTube's recommendation system in spreading controversial content that in the future could include Deepfake fabrications, only one comment explicitly mentions the role of recommendations.[91, 92] Commenting on the Bloomberg video on Deepfake, this user writes: "i'm scared they are recommending me all these videos after i watched the shane dawson video literally videos like this never showed up before wtf."[93]

Users here refer to a popular YouTube conspiracy theories series by Shane Dawson that is also referenced by other commenters across multiple of the five videos analysed in this study. Instead of directly talking about the recommendation system, commenters, for example, ask "Who else is here from Shane?" or "Who else came from Shane's new video?"[94, 95] These comments get notably many likes, sometimes in the hundreds, suggesting that many more users indicate they also got to the video about Deepfake because of a recommendation. By asking who else was recommended to a Deepfake after seeing the Shane Dawson video, users are looking to confirm their own hypotheses about YouTube's recommendation system. Arguably, these commenters are using YouTube as a discursive space to discuss how content circulates on the platform.[96] But while plenty of commenters seem to be aware of the basic workings of YouTube recommendations, the system of recommendations is only referenced, not critiqued.

Instead, most comments that focus on the role of the platform make a point that it should technically be possible to mark or to ban Deepfakes. In this context, one can distinguish between comments that advocate self-regulation and those saying governments should intervene. The most common type of solution that is proposed by users are self-regulatory solutions for YouTube that

[91] Zeynep Tufekci. "YouTube, the Great Radicalizer." *The New York Times*, March 10, 2018, https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html (accessed May 3, 2019).
[92] Bernhard Rieder, Ariadna Matamoros-Fernández, and Òscar Coromina. "From Ranking Algorithms to 'Ranking Cultures': Investigating the Modulation of Visibility in YouTube Search Results." *Convergence*, 24.1 (2018): 50–68.
[93] k e y. "i'm scared they are recommending me all these videos after i watched the shane dawson video literally videos like this never showed up before wtf." February 1, 2019. Comment on Bloomberg, "It's Getting Harder to Spot a Deep Fake Video." https://www.youtube.com/watch?v=gLoI9hAX9dw&lc=UgzWFZ049q-oJH4KoRV4AaABAg (accessed April 22, 2019).
[94] Tori Monahan. "who else is here from Shane 💕." January 31, 2019. Comment on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgxxPZOJQIvfNBaXb2p4AaABAg (accessed April 22, 2019).
[95] Johanna Harak. "Who else came from Shane's new video?" February 1, 2019. Comment on Bloomberg, "It's Getting Harder to Spot a Deep Fake Video." https://www.youtube.com/watch?v=gLoI9hAX9dw&lc=Ugzw1KOlSCvjc2fYXFh4AaABAg (accessed April 22, 2019).
[96] Laura Reed & danah boyd. Who Controls the Public Sphere in an Era of Algorithms? Questions and Assumptions. New York: Data & Society Research Institute, 2016. https://datasociety.net/pubs/ap/QuestionsAssumptions_Background-primer_2016.pdf (accessed February 10, 2019).

aim to validate the authenticity of content.[97] Comments that advocate government regulation often suggest to make the Deepfake technology itself illegal.[98] These commenters are often challenged by others through thread replies that claim that bans can never completely prevent a technology from being used.[99] Another 23% of the comments that can be seen to touch the topic of regulation imply that no type of regulation will be able to stop Deepfakes and the world will need to get used to it.[100]

        One of the things these comments have in common is that they do not focus on individual skills for identifying Deepfakes but on regulatory solutions (or their impossibility). None of these solutions get a lot of support from other users though, as likes are only given sparsely to these comments.

---

[97] Chick Bait. "This will force the use of the blockchain for content creation verification of authenticity." October 30, 2018. Comment on Matthias Niessner, "Face2Face: Real-time Face Capture and Reenactment of RGB Videos (CVPR 2016 Oral)." https://www.youtube.com/watch?v=ohmajJTcpNk&lc=UgwufMD8eF25Uz0o4yp4AaABAg (accessed April 22, 2019).

[98] Example: LunaMoonlight100. "This needs to be banned and destroyed. I can already see this as the golden apple!" April 8, 2016. Comment on Matthias Niessner, "Face2Face: Real-time Face Capture and Reenactment of RGB Videos (CVPR 2016 Oral)." https://www.youtube.com/watch?v=ohmajJTcpNk&lc=UgiKP8nKkRI8oXgCoAEC (accessed April 22, 2019).

[99] rixille. "Laws won't deter many different kinds of people from using this technology. Once a technology is made, it will slip into the hands of someone." March 12, 2019. Comment reply on Pandaman, "Legislature, anybody? Are we going to let this recipe for pure chaos just go ahead out of control? Great." January 25, 2019. https://www.youtube.com/watch?v=gLoI9hAX9dw&lc=UgzwpTdPZCkzmDJzTgt4AaABAg (accessed April 22, 2019).

[100] antigen4. "get used to it. this is our new world." April 7, 2018. Comment reply on Rapunzels Toy Reviews and Super Adorb Adventures!, "This is DANGEROUS." April 4, 2016. https://www.youtube.com/watch?v=ohmajJTcpNk&lc=UgiHoGWWDHYLKngCoAEC.8DTWSw0Oc2P8eiwU6loCOY (accessed April 22, 2019).

5. Discussion & Conclusion: Questioning Digital Literacy

5.1 Putting Comments in Context

This study set out to analyse the discourse on digital literacy in connection to the phenomenon of Deepfakes. The digital literacy skills that are discussed in this context have been found to be effective at identifying the current generation of Deepfakes, which can be falsified through basic audio-visual validation techniques. As to be expected, only a small portion of the top comments on videos Deepfakes are about these techniques, with the most popular group of comments focussing on the technology itself. While only a small group of commenters elaborated on techniques that can unmask Deepfakes as fabrications, such comments received a significant amount of likes compared to the majority of comments that receive no interaction from other users.

Comments that claim that it is easy for everyone to identify Deepfakes receive a considerable amount of counter responses. Such threaded replies are arguably the best manifestation of YouTube as a discursive space, in which commenters react to each other as opposed to only the video. The threaded responses here share a lot of similarities with one of the most popular types of concerns that is shared throughout the comments: Deepfakes are said to become more sophisticated and harder to identify by individuals in the future. After all, even the Zuckerberg video that was found to have been manipulated through basic techniques and announced nothing less than a "world without Facebook" on April Fools' was believed by many of the millions of people that had seen the video.[101,][102] As both scholars and commenters indicate, the basic contents of message can be enough for people to accept something as true, especially when aligns with someone's controversial views or when it inflames.[103] This suggests that in situations in which digital literacy is arguably most desired, the people that are most susceptible to the manipulated message might not even be cued to question the content.

On the contrary, being unduly critical of a piece of content might also lead to unwanted outcomes. While being critical is a key part of being digital literate, this can backfire when commenters see their general mistrust against (mainstream) media sources as being genuinely critical. Instead of critically assessing a questionable message, some commenters simply do not believe certain messengers, period. If individuals are heavily biased for or against or certain media, these biases might

---

[101] Adam Clark Estes. "How One Man Convinced a Staggering Number of People That Mark Zuckerberg Was Deleting Facebook." *Gizmodo,* April 18, 2018. https://gizmodo.com/how-one-man-convinced-a-staggering-number-of-people-Tha-1825323986 (accessed June 4, 2019).

[102] Dan Evon. "Did Mark Zuckerberg Announce That He's Closing Facebook?" *Snopes*, April 18, 2018. https://www.snopes.com/fact-check/mark-zuckerberg-closing-facebook/ (accessed June 3, 2019).

[103] Cass Sunstein. *Republic.com*. Princeton: Princeton University Press, 2001, 65–67.

bleed into people's assessments of other sources. This suggests that the main digital literacy advice of 'being critical' does not necessarily produce the desired outcomes.[104, 105]

5.2 Limitations

Because the comments analysed in this study are unevenly distributed across the five analysed videos — which are not equal in tone and form themselves — certain topics in the discourse analysis might be overrepresented while others are underrepresented. The design of this study also did not allow for a separate analysis of differences between what is underneath the individual videos, which would be interesting to explore in another study. In addition, the singular categorisation of comments did not allow for overlaps and a result, could not fully capture the plurality of what was said. While this is certainly a limitation, assigning multiple categories to single comments arguably would have made it even more difficult to express general claims about the discourse through percentages.

5.3 Conclusion

While digital literacy has proven success in equipping individuals against manipulated content, this study calls into question whether skills at the individual level offer an effective solution for protecting the mass public against ever more advanced forms of misinformation. Digital literacy is put forward by the prevailing online content platforms as a key tool for solving the 'problem of misinformation'.[106] However, as this study shows, both scholars and YouTube commenters debate whether all individuals are equally capable of successfully identifying manipulated messages. The reasons given are multifold but come down to individuals either lacking the required level of digital literacy skills or lacking the right level of cautious, being either unconscionably critical of a content's source, or largely unaware.

In the light of this study's findings, we need to question more critically whether individual solutions can solve a societal problem. Digital literacy has merits but at the same time might be seized as an opportunity for platforms to shift the responsibility for validating content onto its users. This study supports some of boyd's claims that the virtue of being critical towards the media and messages one receives can be "weaponized" to reduce platforms' own responsibilities in this regard.

---

[104] danah boyd. "You Think You Want Media Literacy… Do You?." *Points* (blog), *Data & Society Research Institute,* March 9, 2018. https://points.datasociety.net/you-think-you-want-media-literacy-do-you-7cad6af18ec2 (accessed March 10, 2019).

[105] Example: Gravity False Canada. "BallGod9009 Lol you're an idiot! Buzzfeed is fake news! 😂😂😂." April 19, 2018. Comment reply on deagleninja, "Buzzfeed talking about "trusted news sources".....oh the irony." April 19, 2018. https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=Ugxc815zKKGU8sWH7zR4AaABAg.8fCx_mXzjJ_8fDgehg _8MO (accessed April 22, 2019).

[106] Monica Bulger & Patrick Davison. *The Promises, Challenges, and Futures of Media Literacy.* New York: Data & Society Research Institute, 2018, 5. https://datasociety.net/pubs/oh/DataAndSociety_Media_Literacy_2018.pdf (accessed February 13, 2019).

While YouTube's recommendation system is found to play a significant role in the circulation of radical content, favouring videos about controversial topics such as Deepfakes, commenters are not found to be critical of the video platform itself. Instead of trying to make individuals more conscious and critical of the media content they receive through digital literacy techniques, the digital literacy debate should focus on making the public more literate about the (algorithmic) workings of the platforms that circulate much of this content.

5.4 Further Research

For a follow-up study, it would be interesting to examine how users of digital platforms share knowledge about the technical workings of such platforms. As we have seen, digital platforms provide people with an increasing share of the content they consume. A cross-platform analysis of users' understanding of the platforms algorithmic workings would be an interesting measure of digital literacy. Due to its limited scale, the textual analysis in this study only considered popular Deepfakes of high-profile figures. For further research, it would also be good to consider other genres of Deepfakes including porn-related videos to examine how commenters think of about these videos. This study could focus on how Reddit users share knowledge about Deepfakes as Reddit was one of the first platforms on which the technology was discussed.[107]

---

[107] Mark Wilson. "The War On What's Real." *Fast Company,* June 3, 2018. https://www.fastcompany.com/90162494/the-war-on-whats-real (accessed June 3, 2019).

Bibliography

Allcott, Hunt, Matthew Gentzkow and Chuan Yu. "Trends in the Diffusion of
	Misinformation on Social Media." Working Paper, 2018. https://web.stanford.edu/
	~gentzkow/research/fake-news-trends.pdf (accessed February 19, 2019).

Burgess, Jean & Joshua Green. *YouTube: Online Video and Participatory Culture.* Malden:
	Polity Press, 2009, 69.

boyd, danah. *It's complicated.* New Haven: Yale University Press, 2014.

boyd, danah. "Google and Facebook Can't Just Make Fake News Disappear." *Points* (blog), *Data &
	Society Research Institute*, March 27, 2017. https://points.datasociety.net/google-and-
	facebook-cant-just-make-fake news disappear-48f4b4e5fbe8 (accessed February 10, 2019).

boyd, danah. "You Think You Want Media Literacy… Do You?." *Points* (blog), *Data & Society
	Research Institute*, March 9, 2018. https://points.datasociety.net/you-think-you-
	want-media-literacy-do-you-7cad6af18ec2 (accessed March 10, 2019).

Brabham, Daren C. "THE MYTH OF AMATEUR CROWDS." Information, Communication and
	Society, 15.3 (2012): 394–410.

Bulger, Monica & Patrick Davison. *The Promises, Challenges, and Futures of Media Literacy.*
	New York: Data & Society Research Institute, 2018. https://datasociety.net/pubs/oh/
	DataAndSociety_Media_Literacy_2018.pdf (accessed February 13, 2019).

Caplan, Robyn & danah boyd. *Who Controls the Public Sphere in an Era of Algorithms? Mediation,
	Automation, Power.* New York: Data & Society Research Institute, 2016.
	https://datasociety.net/pubs/ap/MediationAutomationPower_2016.pdf (accessed
	February 10, 2019).

Couldry, Nick, & Joseph Turow. "Advertising, Big Data and the clearance of the public realm:
	Marketers' new approaches to the content subsidy." *International Journal of
	Communication.* 8 (2014): 1710–1726.

Evon, Dan. "Did Mark Zuckerberg Announce That He's Closing Facebook?" *Snopes*, April 18,
	2018. https://www.snopes.com/fact-check/mark-zuckerberg-closing-facebook/ (accessed
	June 3, 2019).

Estas, Adam Clark. "How One Man Convinced a Staggering Number of People That Mark
	Zuckerberg Was Deleting Facebook." Gizmodo, April 18, 2018.
	https://gizmodo.com/how-one-man-convinced-a-staggering-number-of-people-
	tha-1825323986 (accessed June 4, 2019).

Friesem, Yonty. "Teaching Truth, Lies and Accuracy in the Digital Age: Media
	Literacy as Project-Based Learning." *Journalism & Mass Communication Educator.*
	00.0 (2019): 1–14.

Gillespie, Tarleton. "Algorithm." In *Digital Keywords: A Vocabulary of Information Society and Culture*. Ed. by Benjamin Peters. Princeton, NJ: Princeton University Press, 2016, 28–31.

Haysom, Sam. "People are using face-swapping tech to add Nicolas Cage to random movies and what is 2018." Mashable, January 21, 2018. https://mashable.com/2018/01/31/nicolas-cage-face-swapping-deepfakes/?europe=true (accessed June 3, 2019).

Huckin, Thomas. "Critical discourse analysis and the discourse of condescension." *Discourse Studies in Composition*. Edited by E. Barton and G. Stygall. Cresskill, NJ: Hampton Press, 2002, 155–176.

Jenkins, Henry; Sam Ford & Joshua Green. *Spreadable Media: Spreadable media : Creating Value and Meaning in a Networked Culture*. New York: New York University Press, 2013.

Klaas, Brian. "Deepfakes are coming. We're not ready." *The Washington Post*, May 14, 2019. *https://www.washingtonpost.com/opinions/2019/05/14/deepfakes-are-coming-were-not-ready/* (accessed May 18, 2019*).*

Mihailidis, Paul & Samantha Viotty. "Spreadable Spectacle in Digital Culture: Civic Expression, Fake News, and the Role of Media Literacies in "Post-Fact" Society." *American Behavioral Scientist*. 61.4 (2018): 441–454.

Maras, Marie-Helen & Alex Alexandrou. "Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos." *The International Journal of Evidence & Proof*. N.A. (2018): 1–8.

Martin, Allen. "Literacies for the digital age." *Digital Literacies for Learning*. Edited by Alleen Martin, Dan Madigan. London: Facet, 2006, 3–25.

Murthy, Dhiraj & Sanjay Sharma. "Visualizing YouTube's comment space: online hostility as a networked phenomena." *New Media & Society*, 21.1 (2019): 191–213.

Pariser, Eli. *The Filter Bubble: What the Internet is Hiding from You*. New York: Penguin, 2011.

Poell, Thomas & van Dijck, José. "Social media and new protest movements." *The SAGE Handbook of Social Media*. Edited by J. Burgess, A. Marwick, & T. Poell. Los Angeles: SAGE reference, 2018, 546–561.

Reed, Laura & boyd, danah. *Who Controls the Public Sphere in an Era of Algorithms? Questions and Assumptions*. New York: Data & Society Research Institute, 2016. https://datasociety.net/pubs/ap/QuestionsAssumptions_Background-primer_2016.pdf (accessed February 10, 2019).

Rieder, Bernard. "Video Info and Comments Module." *YouTube Data Tools*. N.D. https://tools.digitalmethods.net/netvizz/youtube/mod_video_info.php (accessed April 22, 2019).

Rieder, Bernhard, Ariadna Matamoros-Fernández, and Òscar Coromina. "From Ranking Algorithms to 'Ranking Cultures': Investigating the Modulation of Visibility in YouTube Search Results." *Convergence*, 24.1 (2018): 50–68.

Rotman, Dana, Jennifer Golbeck & Jennifer Preece. "The community is where the rapport is – on sense and structure in the YouTube community." In: *Proceedings of the fourth international conference on Communities and Technologies (C&T '09), University Park, PA, 25–27 June 2009.* New York, NY: ACM, 2009, 41–50.

Solsman, Joan E. "YouTube's AI is the puppet master over most of what you watch." *CNET*, January 10, 2018 (accessed May 30, 2019).

Sunstein, Cass. *Republic.com*. Princeton: Princeton University Press, 2001, 65–67.

Thalen, Mikael (@MikaelThalen). "I've gone down a black hole of the latest DeepFakes and this mashup of Steve Buscemi and Jennifer Lawrence is a sight to behold." Tweet, January 29, 2019. https://twitter.com/MikaelThalen/status/109034993226609 4593/video/1 (accessed February 10, 2019).

Thelwall, Mike, Pardeep Sud & Farida Vis. "Commenting on YouTube videos: From guatemalan rock to El Big Bang." *Journal of the Association for Information Science and Technology*. 63.3 (2012): 616–629.

Tufekci, Zeynep. "YouTube, the Great Radicalizer." *The New York Times*, March 10, 2018. https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html (accessed May 3, 2019).

Van Dijck, José. *The Culture of Connectivity.* Oxford: Oxford University Press, 2013.

Van Dijck, José, Thomas Poell & Martijn de Waal. *The Platform Society: Public Values in a Connective World.* New York: Oxford University Press, 2018, 21.

Van Es, Karin. "YouTube's Operational Logic: "The View as Pervasive Category." *Television and New Media.* N.A. (2019): 1–17.

VPRO Tegenlicht. "Deep Fake News." *VPRO Tegenlicht*, VPRO, 18 November 2018. https://www.vpro.nl/programmas/tegenlicht/kijk/afleveringen/2018-2019/ deep-fake-news.html (accessed March 10, 2019).

Walker, Anna Shoenfeld. "Preparing Students for the Fight Against False Information With Visual Verification and Open Source Reporting Journalism & Mass Communication Educator." *Journalism & Mass Communication Educator.* N.A. (2019): 1–13.

Konrad Weber. "8 steps to verify Deep Fake videos." *Medium* (blog), *Konrad Weber,* November 8, 2018. https://medium.com/@konradweber/8-steps-to-verify-deep-fake-videos-1dfc408568c4 (accessed May 25, 2019).

Wilson, Mark. "The War On What's Real." *Fast Company*, June 3, 2018. https://www.fastcompany.com/90162494/the-war-on-whats-real (accessed June 3, 2019).

Other Sources

Video Works

Bayview Drive Films. "A WORLD WITHOUT FACEBOOK." YouTube video, 3:23, April 1, 2018. https://www.youtube.com/watch?v=ADCPqE4ShGQ#action=share (accessed April 22, 2019).

Bloomberg. "It's Getting Harder to Spot a Deep Fake Video." YouTube video, 2:58, September 27, 2018. https://www.youtube.com/watch?v=gLoI9hAX9dw&t=3s (accessed April 22, 2019)).

BuzzFeedVideo. "You Won't Believe What Obama Says In This Video! 😉." Youtube video, 1:12, Apr 17, 2018. https://www.youtube.com/watch?v=cQ54GDm1eL0 (accessed April 22, 2019).

derpfakes. 2018. "Putin | Face Replacement." YouTube video, 0:37, Feb 26, 2018. https://www.youtube.com/watch?v=hKxFqxCaQcM&feature=youtu.be (accessed April 22, 2019).

Niessner, Matthias. "Face2Face: Real-time Face Capture and Reenactment of RGB Videos (CVPR 2016 Oral)." YouTube video, 6:36, March 17, 2016. https://www.youtube.com/watch?v=ohmajJTcpNk (accessed April 22, 2019).

Video of The Internet. "A WORLD WITHOUT FACEBOOK." Facebook video, 3:23, April 23, 2018. https://www.facebook.com/VIDEOOTI/videos/352738545218184/ (accessed June 3, 2019).

YouTube Comments

ïï Clorox Bleach ïï. "Guys this is a joke and edited, I did research and check on his lips.. As you can see it is not his lips but someone else's edited there." April 13, 2018. Comment on Bayview Drive Films. "A WORLD WITHOUT FACEBOOK." https://www.youtube.com/watch?v=ADCPqE4ShGQ&lc=UgzXlUxKktlrnQI1-jF4AaABAg (accessed April 22, 2019).

antigen4. "get used to it. this is our new world." April 7, 2018. Comment reply on Rapunzels Toy Reviews and Super Adorb Adventures!, "This is DANGEROUS." YouTube comment, April 4, 2016. https://www.youtube.com/watch?v=ohmajJTcpNk&lc=UgiHoG WWDHYLKngCoAEC (accessed April 22, 2019).

Casey, Colleen. "Perhaps instead of 'insulting' people who do not dedicate their lives to 'technology' and non-living things, and choose to live in the 'real world of Human Beings'...maybe help to explain to those of us who do not have those 'skill sets' understand could be a bit more helpful. I admit, I'm not tech literate, and do depend on those willing and caring enough to explain it to me in a way I can understand in order to navigate what does appear to look 'very real'. I have great respect for those who work in the Tech fields, I just am not that type of 'thinker' and have a different set of skills...doesn't make either or better or worse. Thanks." Comment reply on ThePromisedWLAN. "This is frightening technology for media manipulation," October 7, 2016. https://www.youtube.com/watch?v=ohmajJTcpN k&lc=UgjJzYe0lhU_JHgCoAEC (accessed April 22, 2019).

Chick Bait. "This will force the use of the blockchain for content creation verification of authenticity." October 30, 2018. Comment on Matthias Niessner, "Face2Face: Real-time Face Capture and Reenactment of RGB Videos (CVPR 2016 Oral)." https://www.youtube.com/watch?v=ohmajJTcpNk&lc=UgwufMD8eF25Uz0o4yp4AaAB Ag (accessed April 22, 2019).

deagleninja, "Buzzfeed talking about "trusted news sources".....oh the irony." April 19, 2018. Comment on Comment on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc= Ugxc815zKKGU8sWH7zR4AaABAg (accessed April 22, 2019).

Diaz, Aaron. "Lmao rely on trusted news sources, is not buzzfeed the most distrusted 😂😂😂." April 19, 2018. Comment on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1 eL0&lc=UgwSOpLvRgsXRZpIRJh4AaABAg (accessed April 22, 2019).

Gravity False Canada. "BallGod9009 Lol you're an idiot! Buzzfeed is fake news! 😂😂😂." April 19, 2018. Comment reply on deagleninja, "Buzzfeed talking about "trusted news sources".....oh the irony." April 19, 2018. https://www.youtube.com/watch?v=cQ54GDm1eL0&lc= Ugxc815zKKGU8sWH7zR4AaABAg.8fCx_mXzjJ_8fDgehg_8MO (accessed April 22, 2019).

Greene, William. "It look like Obama got beat up... a swollen left jaw and perhaps maybe a black eye with some make-up covering it up...damn!." December 14, 2018. Comment on BuzzFeedVideo. "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=Ugzvegk2gkebr6yd GQx4AaABAg (accessed April 22, 2019.

Harrison, G.L. "I knew it wasn't him talking straight off the bat." Comment on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉."
https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgwHsaBv1QuCLBjgYpB4Aa ABAg (accessed April 22, 2019).

Harak, Johanna. "Who else came from Shane's new video?" February 1, 2019. Comment on Bloomberg, "It's Getting Harder to Spot a Deep Fake Video."
https://www.youtube.com/watch?v=gLoI9hAX9dw&lc=Ugzw1KOlSCvjc2fY XFh4AaABAg (accessed April 22, 2019).

Hutchins, Owen. "Because buzzfeed is a trusted news source loooool." April 19, 2018. Comment on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉."
https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgyBF6R1UuTdG1OpUAZ4 AaABAg (accessed April 22, 2019).

Ibieiniid. "all it would take to convince some people is for them to WANT to believe what theyre seeing. If there was a grainy video of Obama saying 'I murdered Seth Rich after a pedo party at Comet Ping Pong', whats to say some gullible idiot wouldnt be planning Pizzagate part 2 immediately? Truth is, there ARE some people you can't help, because you're an internet nobody (not an insult, just fact). Jordan Peele has the stage to try to help those people." April 20, 2018. Comment reply on moiraine_damodred, "if you couldn't immediately tell that was computer generated I can't help you." April 19, 2018.
https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgyO2NJmqvH4JL861OF4Aa ABAg.8fDkcj8tmPb8fGLmUCU_vt (accessed April 22, 2019).

Isaac, Michael. "Noticed his face especially around mouth is swollen and the voice isn't quite right. Great observation!" April 18, 2018. Comment reply on massessum1, "You can see, if you look closely, that Obama's face is animated, especially his smile lines. Plus, the voice is not quite right. Pretty damn good, though." YouTube comment, April 18, 2018.
https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=Ugxnp-hbnr7yMsO9Yap4AaABAg.8fBZ3O89cHG8fBgDFtQIbE (accessed April 22, 2019).

Kairas_love. "Eve[r]yone on here says it looks fake because you know that from the start. The issue being adressed here is that the mass public wouldn't be able to tell. Say this was posted on television and your parent[s] or Grandparents saw this. Would they know it was fake? Probably not. Our brains don[']t need too much detail to believe in something. He is animated so well that a quick glace at the screen and hearing the familiar voice sound would make anyone believe in it enough to cause conflict." April 17, 2018. Comment on BuzzFeedVideo. "You Won't Believe What Obama Says In This Video! 😉."
https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgxaLrzhpip9TjJo0ut4AaABA A (accessed April 22, 2019).

k e y. "i'm scared they are recommending me all these videos after i watched the shane dawson video literally videos like this never showed up before wtf." February 1, 2019. Comment on Bloomberg, "It's Getting Harder to Spot a Deep Fake Video." https://www.youtube.com/watch?v=gLoI9hAX9dw&lc=UgzWFZ049q-oJH4KoRV4AaABAg (accessed April 22, 2019).

largelysubatomic. "this isn't even a good example of how effective this technology is right now let alone in the future. this is just a meme vid from buzzfeed while nation states, corporations, and other various non-state actors are already capable of producing something far more convincing." April 18, 2018. Comment reply on Mauri QHD, "if you are not 100% sure is not him as soon as 0:02, you are part of the problem." April 18, 2018. https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgwdX1dvbY7lZcCrEL14AaABAg.8f8p_CQcL5g8fAAsDuTyOi (accessed April 22, 2019).

Leadsticks. "0:51 This is all fake. The eyes don't blink at the same time. Still don't believe me? Look at the eyebrows." January 28, 2019. Comment on Bloomberg, "It's Getting Harder to Spot a Deep Fake Video." https://www.youtube.com/watch?v=gLoI9hAX9dw&lc=UgxeXgci5R9wc3KhP-R4AaABAg (accessed April 22, 2019).

LunaMoonlight100. "This needs to be banned and destroyed. I can already see this as the golden apple!" April 8, 2016. Comment on Matthias Niessner, "Face2Face: Real-time Face Capture and Reenactment of RGB Videos (CVPR 2016 Oral)." https://www.youtube.com/watch?v=ohmajJTcpNk&lc=UgiKP8nKkRI8oXgCoAEC (accessed April 22, 2019).

massessum1. "You can see, if you look closely, that Obama's face is animated, especially his smile lines. Plus, the voice is not quite right. Pretty damn good, though." April 18, 2018. Comment on on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=Ugxnp-hbnr7yMsO9Yap4AaABAg (accessed April 22, 2019).

Mauri QHD. "if you are not 100% sure is not him as soon as 0:02, you are part of the problem." April 18, 2018. Comment on BuzzFeedVideo, "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=UgwdX1dvbY7lZcCrEL14AaABAg.8f8p_CQcL5g8fAAsDuTyOi (accessed April 22, 2019).

Meme mine. "Too funny! Bloomberg will be the first ones to use this new technology to spread their fake news." October 17, 2018. Comment on Bloomberg, "Bloomberg. "It's Getting Harder to Spot a Deep Fake Video." https://www.youtube.com/watch?v=gLoI9hAX9dw&lc=UgxQC3UX77HDsm6Ut3B4AaABAg (accessed April 22, 2019).

moiraine_damodred, "if you couldn't immediately tell that was computer generated I can't help

        you." April 19, 2018. Comment on BuzzFeedVideo, "You Won't Believe What Obama

        Says In This Video! 😉." https://www.youtube.com/watch?v=cQ54GDm1eL0&lc=

        UgyO2NJmqvH4JL861OF4AaABAg (accessed April 22, 2019).

Monahan, Tori. "who else is here from Shane 💕." January 31, 2019. Comment on BuzzFeedVideo,

        "You Won't Believe What Obama Says In This Video! 😉." https://www.youtube.com/

        watch?v=cQ54GDm1eL0&lc=UgxxPZOJQIvfNBaXb2p4AaABAg (accessed April 22,

        2019).

Pandaman, "Legislature, anybody? Are we going to let this recipe for pure chaos just go ahead out

        of control? Great." January 25, 2019. Comment on Bloomberg, "It's Getting Harder to

        Spot a Deep Fake Video." https://www.youtube.com/watch?v=gLoI9hAX9dw&lc=

        UgzwpTdPZCkzmDJzTgt4AaABAg (accessed April 22, 2019).

PBS NewsHour. "WATCH: Facebook's Mark Zuckerberg to cooperate on Russian investigation

        after 2016 election." YouTube video, 8:31, September 21, 2017.

        https://www.youtube.com/watch?v=wsUhyiO31DY (accessed June 3, 2019).

PUMPA DOUR. "I have never ever believed the mass media before, so this shit is not too scary to

        me. I will continue not believing anything i see and hear." April 10, 2016. Comment on

        Matthias Niesner, "Face2Face: Real-time Face Capture and Reenactment of RGB Videos

        (CVPR 2016 Oral)." https://www.youtube.com/watch?v=ohmajJTcpNk&lc=Ugg

        2AmnPj8qDcHgCoAEC (accessed April 22, 2019).

Rapunzels Toy Reviews and Super Adorb Adventures!. "This is DANGEROUS." April 4, 2016.

        Comment on Matthias Niessner, "Face2Face: Real-time Face Capture and Reenactment of

        RGB Videos (CVPR 2016 Oral)." https://www.youtube.com/watch?v=ohmajJTcp

        Nk&lc=UgiHoGWWDHYLKngCoAEC (accessed April 22, 2019).

rixille. "Laws won't deter many different kinds of people from using this technology. Once a

        technology is made, it will slip into the hands of someone." March 12, 2019. Comment

        reply on Pandaman, "Legislature, anybody? Are we going to let this recipe for pure chaos

        just go ahead out of control? Great." YouTube comment, January 25, 2019.

        https://www.youtube.com/watch?v=gLoI9hAX9dw&lc=UgzwpTdPZCkzmDJzTgt4AaA

        BAg (accessed April 22, 2019).

Schmeller, Fanny. "just wait until adobe release their photoshop for audio software that theyv've

        been demoing. we wont have a clue whats real." November 27, 2016. Comment reply on

        ThePromisedWLAN, "This is frightening technology for media manipulation," October 7,

        2016. https://www.youtube.com/watch?v=ohmajJTcpNk&lc=UgjJzYe0lhU_JHgCoAEC

        (accessed April 22, 2019).

The Obama White House. "President Obama Speaks to the Open Government Partnership Global
    Summit." YouTube video, 2:09, December 17, 2016. https://www.youtube.com/watch?v=
    -mIaZPqoOYw (accessed June 13, 2019).

ThePromisedWLAN. "This is frightening technology for media manipulation," October 7, 2016.
    Comment on "Matthias Niessner. "Face2Face: Real-time Face Capture and Reenactment
    of RGB Videos (CVPR 2016 Oral)." https://www.youtube.com/watch?v=ohmajJTc
    pNk&lc=UgjJzYe0lhU_JHgCoAEC (accessed April 22, 2019).

List of Figures

Figure. 1. Comparison of the Obama Deepfake on the left and the original shot on the right.

BuzzFeedVideo. "You Won't Believe What Obama Says In This Video! 😉." YouTube video, 1:12,
    April 17, 2018. https://www.youtube.com/watch?v=cQ54GDm1eL0 (accessed April 22,
    2019).

The Obama White House. "President Obama Speaks to the Open Government Partnership Global
    Summit." YouTube video, 2:09, December 17, 2016. https://www.youtube.com/watch?v
    =-mIaZPqoOYw (accessed June 13, 2019).

Figure 2. Screenshot of the Putin video with *SNL*-actor Beck Bennett on the left and on the right
a version where Bennett's face is swapped for Putin's.

derpfakes. 2018. "Putin | Face Replacement." YouTube video, 0:37, Feb 26, 2018.
    https://www.youtube.com/watch?v=hKxFqxCaQcM&feature=youtu.be (accessed April
    22, 2019).

Figure 3. Comparison of the manipulated Zuckerberg on the left and the original footage
on the right.

Bayview Drive Films. "A WORLD WITHOUT FACEBOOK." YouTube video, 3:23, April 1,
    2018. https://www.youtube.com/watch?v=ADCPqE4ShGQ#action=share (accessed
    April 22, 2019).

PBS NewsHour. "WATCH: Facebook's Mark Zuckerberg to cooperate on Russian investigation
    after 2016 election." YouTube video, 8:31, September 21, 2017.
    https://www.youtube.com/watch?v=wsUhyiO31DY (accessed June 3, 2019).

Images: Composition of YouTube Screencaptures