Trust and Accountability in a Digital Age

ONORA O'NEILL

1. Remembering Stewart Sutherland

I have a very particular reason to be grateful to Stewart Sutherland, our late President, which is connected to some of the themes of this lecture, so want to begin by recalling a long conversation I had with him on these topics.

He rang me up in 2000 to ask whether I could give the Gifford Lectures in Edinburgh in 2001 on issues in bioethics, and suggested autonomy as the theme. I replied (rather peevishly, I now think) that I was fed up with the many discussions of autonomy that were swirling around, especially in bioethics. They seemed to me to concentrate on individual choice and self-expression, but to have limited moral weight. I had become irritated with those discussions in part because I had long worked on Kant's practical philosophy and had concluded that contemporary discussions of autonomy often claimed a Kantian pedigree, but repeatedly and inaccurately ascribed individualistic and morally shallow conceptions of autonomy to him. They were free-riding on Kant's reputation, travestying his conception of autonomy, and had lost touch with the weightier account of autonomy that Kant had offered.

As I saw it, and as I later managed to articulate, there is a vast gulf between contemporary conceptions of individual autonomy and Kant's conception of autonomy, which was his term of art for principles that could be principles for all. There is in my view zero prospect of finding a coherent reading of Kant's *Groundwork*, in which he claims that there are five different, complex but equivalent formulations of the Categorical Imperative (of which only one uses the term 'autonomy') if one imagines that Kantian autonomy is some form of individual autonomy.

Stewart listened with care, as he always did, and we talked for quite a long time, finally agreeing that I would give the lectures, but that I would focus not on *individual autonomy* as currently understood, which had taken root in the bioethical literature, but on differences between those accounts and the conception of *principled autonomy*

doi:10.1017/S0031819119000457 © The Royal Institute of Philosophy, 2019
First published online 14 October 2019

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

that can be found in Kant's writings, and the more robust approaches to ethical justification for which it provides an opening.

So I gave the 2001 Gifford Lectures, sorted out my thinking about differences between individual autonomy and Kantian autonomy, and in the process realised that there was a good deal of interesting history to be unearthed. In particular I realised that that the very word 'autonomy' was still used in its classical jurisprudential sense, that stressed laws and law-like norms rather than individual willing, long after Kant's day. For example, the only use of the term 'autonomy' that I could find in John Stuart Mill's work, was in *On Representative Government*, where he characterises autonomous cities as cities that make their own laws, contrasting them with colonies that have their laws given to them by their mother cities. Mill had plenty to say about individuality: but he did not conflate it with autonomy, which he too understands in its classical quasi-jurisprudential sense.

So I am profoundly grateful to Stewart for that conversation, which shaped my Gifford Lectures and the book that grew out of them.[1] And much more. But turning away from individualistic assumptions has deep implications for normative inquiry, and in particular for normative inquiry that focuses on public rather than on private goods, hence on the standards that just social and cultural institutions and democratic polities must meet. As I came to see, a focus on the classical view of autonomy also makes it possible to articulate connections between autonomy and trust, and to offer a clearer account of the fundamental importance of trustworthiness.

2. Trustworthiness before Trust

Individualistic conceptions of autonomy offer a narrow lens, and make little contribution to placing or refusing trust well. Yet placing trust intelligently matters hugely for communication, for public reasoning, for justice and for democracy.

Swathes of contemporary discussion approach trust as if it were no more than a *generic attitude* that can be revealed by asking people, as pollsters do, what and whom they trust or how much they trust different types of product, office-holder and institutions. Opinion polls can provide fascinating information about generic attitudes of trust and mistrust. They have repeatedly shown, for example, that

[1] Onora O'Neill, *Autonomy and Trust in Bioethics* (Cambridge: Cambridge University Press, 2002).

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

in the UK politicians and journalists are not much trusted, whereas judges and nurses are highly trusted (for what it is worth, scientists and academics usually rank high on these popular subjective metrics).

Polls provide reputational evidence, but not the evidence most needed to place and refuse trust well.[2] Placing trust well needs more than reputational evidence, since reputations can be undeserved. In placing trust well we aim to align it with *trustworthiness,* and to avoid placing it in *untrustworthy* products, individuals or institutions. There is little to be said in favour of placing 'more trust' in anything or anyone unless we can distinguish well-directed from misdirected trust. Trustworthiness matters more than trust.

However, when we aim not simply to 'increase trust' or 'restore trust', but to place and refuse it well, evidence about typical attitudes provides too little. Placing and refusing trust with discrimination is hard. It needs judgement and information, evidence and care. We live in a complex world, have connections and relations with others with whom we have no direct contact, and often lack time or access to the sorts of evidence on which we rely in everyday and face-to-face contexts. We may lack *both* evidence that depends on context or culture, gesture or tone, *and* evidence that requires, time, expertise or systematic investigation. Placing or refusing trust in unfamiliar situations with discrimination is therefore taxing, and can be overtaxing.

The problem is ancient. Plato discussed it in *Phaedrus,* where he explains why Socrates did not trust writing, which was then a newish technology that many could not decipher. Because he thought that writing was a deceptive way of communicating, Socrates relied entirely on the spoken word. Plato reports that Socrates was concerned that

> 'When it has once been written down, every discourse roams about everywhere, reaching indiscriminately those with understanding no less than those who have no business with it, and it doesn't know to whom it should speak and to whom it should not. And when it is faulted and attacked unfairly, it always needs its father's [i.e. its author's] support; alone, it can neither defend itself nor come to its own support'.[3]

[2] Consumer rankings of products can be useful for placing trust if, and only if, they are based on responses by a *representative* sample of respondents to *standardised* products. When these conditions are not met, polls cannot provide evidence for trustworthiness.

[3] Plato, *Phaedrus* 275d-e, tr. W. Hamilton, (Harmondsworth: Penguin Books, 1973).

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

The problem, stripped to its bare bones, is that written words can be separated from their authors and contexts–that *speech content* can be separated from *speech act* –and that readers lack information that would be conveyed by tone, gesture and context in face-to-face speech with familiar others, so may not be able to judge whether written words are trustworthy. Readers often need to know something about sources, evidence and transmission, about context and culture, in order to judge whether written words are true or false, plausible or far-fetched, evidence-based or imagined, hostile or helpful.

3. The Old Intermediaries

What Socrates said about the written word raises problems for other types of mediated speech. We have long lived in a world in which a great deal of the information (not to mention misinformation and disinformation) that we encounter reaches us indirectly. A great deal of mediated speech is written, and today most of the written material we encounter is also technologically mediated by print, images or recordings, by film, television and radio, and by digital technologies.

Across a long history, capacities for judging mediated speech acts have been supported by legal, regulatory and cultural measures that make it easier to understand what is meant, and to judge its trustworthiness or untrustworthiness, its truth or falsehood. Typically these measures place requirements on *intermediaries*, who are required to support trustworthy communication, and hold those whose communication proves untrustworthy to account.

Many of the measures on which we now rely were introduced when the invention of printing extended opportunities for distant and potentially anonymous communication, making it harder to judge the trustworthiness of the written word. Printed material was made more assessable by defining the distinctive roles and responsibilities of authors and editors, of publishers and printers, and of other intermediaries, and later by professional standards for journalists, librarians and many others. These intermediaries made it easier for readers and others to identify authors and trace provenance, to seek redress for defamation, fraud and other failings, and to protect intellectual property, including copyright. Legal, regulatory and professional standards were supplemented by cultural conventions for differing genres of writing. Taken together, these innovations made it more feasible to distinguish reporting from propaganda, fact from fiction, honest dealing from fraud, and so on, thereby supporting

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

judgements of trustworthiness and untrustworthiness even when direct contact with originators was lacking. Broadcasting and film were later regulated and organised in analogous ways by defining and distinguishing the roles and responsibilities of originators and intermediaries and ensuring that they could be identified, thereby making it easier to assess mediated speech and images and providing for some redress for untrustworthy claims. Placing (and in some cases enforcing) requirements on originators and intermediaries gave readers, listeners and viewers ways of assessing printed, broadcast and visual material. However, these traditional ways of judging the written word are being disrupted in the digital age.[4]

Of course, these systems of accountability have always been imperfect. They have been criticised for being too complex, too open to corruption and conflicts of interest, for focusing on the wrong drivers and issues, for creating perverse and distracting incentives, and for being too readily gamed. Attempts to marshal incentives for good behaviour by regulation, or by requiring transparency, have often had patchy success, as have approaches that depend on metrics.[5] But some of the problems with digital communication are different.

4. The New Intermediaries

Digital technologies are distinctive in several ways. They seemingly enable *disintermediation* by making it easy, all too easy, to disseminate content widely, rapidly and cheaply, to deliver it to distant others,

[4] The literature is recent, large and growing. See Franklin Foer, *World without Mind: The Existential Threat of Big Tech* (London: Jonathan Cape, 2017); Alan Rusbridger, *Breaking News: the Remaking of Journalism and Why it Matters Now* (Edinburgh: Canongate Books, 2018); Martin Moore, *Democracy Hacked*: *Political Turmoil and Information Warfare in the Digital Age*, (London: Bloomsbury, 2018); Jamie Bartlett, *The People v Tech: How the Internet is killing democracy (and how we save it)* (Penguin Books, 2018); James Williams, *Stand Out of Our Light: Freedom and Resistance in the Attention Economy*, (Cambridge: Cambridge University Press, 2018); Department for Culture, Media and Sport, *Overview of recent dynamics in the UK press market*, (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/720400/180621_Mediatique, April 2018)

[5] For example, anyone familiar with contemporary trends in school and university education will be aware of controversies about approaches to accountability that measure and rank research and teaching and construct league tables.

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

and to by-pass traditional gate-keepers. Many thought initially that they would prove a democratising innovation and would eliminate or reduce privilege and exclusion.

That was (I assume) the hope behind Mark Zuckerberg's famously provocative injunction to 'move fast and break things.' I take it that his thought was that the things that would get broken were useless, even harmful, for example because they imposed unjustifiable constraints and exclusions, or because they ignored or damaged the interests and rights of some citizens or consumers. If that had been the whole story, disintermediation might have offered nothing but gain.

But it was not the whole story. The other side of the story is that the old intermediaries were replaced with new intermediaries, whose contributions are harder to grasp, and who are not and cannot be disciplined by the measures used to discipline the old intermediaries. In consequence it can be hard to judge the source or authority of digitally mediated speech acts, and hard to place and refuse trust in digitally mediated communication intelligently.

Not everything that digital technologies broke or bypassed was without value. The new technologies were organised in ways that often made it harder to identify originators, promoters or intermediaries, hence harder to regulate their action, and harder to place or refuse trust with discrimination. The title of Jonathan Taplin's 2017 book summarises this point: *Move Fast and Break Things: How Facebook, Google and Amazon Cornered Culture and Undermined Democracy*.[6] Cultures have been 'cornered' by eliminating the contributions of many traditional intermediaries, including writers and reporters, editors and publishers, newspapers and broadcasters, in large part by diverting the revenue and the readership that once supported them, to companies that sell online advertisements.[7] Democracy has been 'undermined' by enabling anonymised targeting of citizens by interested parties of many sorts.

At their best, the traditional intermediaries managed to discipline and curate, to fact-check and authenticate, a good deal of the content

[6] Jonathan Taplin, *Move Fast and Break Things: How Facebook, Google and Amazon Have Cornered Culture and Undermined Democracy* (Macmillan, 2018).

[7] See *The Cairncross Review: a sustainable future for journalism, February 2019,* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/779882/021919_DCMS_Cairncross_Review_.pdf; see also Taplin *op. cit* .on the effects of digital technologies on musicians and writers and information compiled by *Authors' Copyright and Licensing Society* on the decline in professional writers' earnings.

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

that members of the public read, heard and viewed, so supported abilities to judge others' claims and action, by providing or requiring a context of assessable evidence and argument that sustained standards and practices of public reason and its cultural and political benefits.[8] Democracy depended in many ways on the work of these traditional intermediaries.

Initially many commentators hoped that digital technologies would spread information more widely, thereby allowing for wider participation and access to information and evidence, and supporting extending and deepening democracy. And so it has in some cases. But it has also supported deception and disinformation, and made it possible to target individuals with misleading or manipulated content that is anonymously orchestrated. Dubious content of many sorts, including what we have come to think of oxymoronically as 'fake news', swills around, making judgement of truth, trustworthiness and untrustworthiness not easier but harder in many contexts.[9]

This I think is no longer disputed, and there are many current discussions of these themes. Notably, in September 2018, Sir Tim Berners-Lee, inventor of the World Wide Web, stated quite explicitly that his early hopes for the World Wide Web that he had initiated have been dashed. He wrote

> 'I've always believed the web is for everyone. That's why I and others fight fiercely to protect it. The changes we've managed to bring have created a better and more connected world. But for all the good we've achieved, the web has evolved into an engine of inequity and division; swayed by powerful forces that use it for their own agendas…'.[10]

5. Accountability On-Line and Off-Line

These problems are often ascribed to the misuse of social media by individuals who abuse, bully, defraud or seduce others, or harm them in other ways. Some individuals do indeed use social media to

[8] See works cited in note 4.

[9] House of Commons Digital, Culture, Media and Sport Committee, *Disinformation and 'fake news': Final Report*, 2019

[10] See Tim Berners-Lee, https://inrupt.com/blog/one-small-step-for-the-web Sept 2018. He has since launched *The Contract for the Web* described as 'an initiative to bring governments, companies, and civil society and web users together to build a roadmap for how we build a web that serves humanity and is a public good for everyone, everywhere'.

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

act reprehensibly and inflict harms (in this case *private harms*) on individuals. However, digital technologies are also used to inflict *public harms* on culture and democracy, and these cannot be laid at the door of individuals who misuse social media.

As I see it, public harms to democracy, to culture, and to serious reporting, are produced not by users of social media, but by the business model of the companies, that control and market opportunities to use data provided by users of social media, and others, in order to market opportunities to inform and misinform, and above all to influence. Moreover, since those using social media are not charged, they are not customers of their service providers. In consequence consumer protection approaches to regulation are not applicable to the provision of targeted content. Individual users of social media provide their data *without payment*, and receive content selected by the actual customers of the digital companies, who remain anonymous.

In short, the business model of the digital providers offers ways of influencing public life, democratic debate and shared cultures by covert means. At present we are only at the start of debates about limiting and redressing these public harms. Some suggest that what is needed is a legally binding 'duty of care' on providers of digital services.[11] This might indeed limit and deter some who distribute content that leads to private harms, where perpetrators can be identified; but it is unlikely to do much to deter public harms, where causal connections are typically unclear.

How then can the online communication be held to account? How can we make it feasible for people to place and refuse trust in online speech with discrimination? The question is urgent, but difficult to answer. There are many current examples of bogus and unrealistic online accountability, hence many areas that currently escape accountability.

Each of us will have some favourite examples, so I offer one or two of mine. One example is the proliferation of hyper-complex online consent requirements in many areas of life; they are particularly striking in medical and commercial life. Patients and consumers are asked to 'consent' to highly complex arrangements, with implications that they are unlikely to grasp. Because the 'consent' forms are long and hard to follow, what passes for consent usually merely requires ticking and clicking. One attempt to show up these bogus consent

[11] See *Online Harms White Paper*, April 2019, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793360/Online_Harms_White_Paper.pdf.

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

practices inserted a clause into a EULA (an *End User License Agreement*) by which those who (supposedly) consented agreed to give their first born to the devil. Of course, many who were fed up with these hyper-complex 'consent' procedures clicked and ticked without reading the EULA, and (supposedly) agreed to this arrangement! As I see it 'clicking and ticking' is not a form of consenting. But the fact that consent forms of absurd length and complexity are in daily commercial, public and professional use shows not that all is well, but rather that consent procedures are no longer a reputable way of securing accountability in such contexts. (There are, of course, worse failings, such as the violation of explicit privacy undertakings in the Cambridge Analytica scandal[12]).

Problems of a different sort arise where the fragmentation of roles and responsibilities held by numerous on-line intermediaries means that none of them is accountable for certain harms. The 'platformisation' of the online world[13] makes it possible and cheap to multiply intermediaries, without ensuring accountability for what they collectively achieve. A familiar example is the avoidance of employer's obligations by platforms that provide taxi services, where drivers are deemed to be self-employed *users* of the platforms, so do not qualify for employee benefits.[14]

6. The Claimant Perspective

However, there are other obstacles to accountability that have nothing to do with digital technologies, or with the structure of providers of digital services. Accountability has also been made harder by fundamental changes in approaches to ethical and political norms during the last century.

During the twentieth century there was a marked narrowing of the range of norms seen as relevant to the public domain, and a particularly marked narrowing of the range of norms that bear on speech and communication. It became less common to focus on requirements or duties, and more common to focus on entitlements or rights. Ethical and political discussion that had once focussed on the agent's

[12] Cf. Note 9.

[13] Nick Srnicek, *Platform Capitalism*, (Cambridge: Polity Press, 2017) esp. ch 2.

[14] And this is not the most troubling case. Since taxi services are licensed by local authorities, there are other routes to their regulation, such as withholding a license to operate in specified areas.

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

question, 'what ought we do?' increasingly addressed the recipient's question 'what ought each of us receive?' or 'what rights do I have?'.

The momentous shift from accounts of duties to accounts of rights has many sources. One is surely the (understandable) discrediting of the narrow versions of the ethics of duty, centered on excessive conceptions of patriotic duty, that became prominent before and during the First World War.[15] Between the wars scepticism about duty received support, not only from the excesses of logical positivism, but from a wide range of writers who criticised and rejected conceptions of duty. In the 1940s this sceptical trend was stemmed, *but only in part*, by seeking international agreement on a narrower range of norms and standards for human rights that were incorporated into the canonical human rights documents.[16]

Rights are indeed fundamental for justice, but they are not all that matters, nor even all that matters for justice. Accounts of justice that prioritise the perspective of claimants or right-holders ignore many of the norms and standards that had once formed part of the ethics of duty, which become invisible once rights were taken as fundamental.

Theories of rights have nothing to say about *imperfect duties,* which leave agents discretion whether to act in particular cases (e.g. generosity and solidarity), since these lack counterpart rights. They are also silent about those *perfect duties* that permit no exceptions but are not duties of justice, since they cannot be adequately enforced by public authorities (e.g. truthfulness and honesty). And they are also silent about further (presumed) *duties of justice that cannot be claimed as rights* because no claimants are specified (e.g. duties to future generations, duties of fairness).[17] Moreover, many human rights make claims that are indeterminate unless and until institutional action identifies counterpart duty-bearers and defines their duties. For all

[15] I have discussed these themes in 'Liberal Justice: Kant, Rawls and Human Rights', *Kantian Review* (2018), 23, 641-59, and 'Ethical and Political Justification in the Twentieth century', the Berggruen Lecture, *Proceedings and addresses of the American Philosophical Association* (2018) 92, 296-309,. Also 'Justice Without Ethics', in John Tasioulas, ed., *Cambridge Companion to the Philosophy of Law* (Cambridge: Cambridge University Press, forthcoming).

[16] *Universal Declaration of Human Rights* (1948); *European Convention on Human Rights* (1950).

[17] John Rawls is a notable exception: he saw justice as combining *two* distinct fundamental principles, the first a principle of equal rights and the second a principle constraining distributions to secure a form of fairness. Hence 'Justice as Fairness'.

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

of these reasons, a range of duties that had long been taken seriously became invisible once rights were taken as foundational.

This narrowing of ethical perspectives has been particularly evident in discussions of norms and standards for communication. Norms that are fundamental for communication include honesty and truthfulness, civility[18] and decency, sincerity and accuracy: all become invisible or marginal once rights are taken as fundamental. Yet these had traditionally been seen as matters of *perfect* duty, which the exercise of personal choice or discretion could not set aside.[19]

Only the two canonical human rights that bear on speech – *freedom of expression* and the *right to privacy*– are now generally taken as fundamental requirements of justice that bear on all speech, and so on all communication. The narrowing of perspectives that arises when rights are treated as fundamental has marginalised other norms that bear on speech, treating them either as a matter of personal choice, or as *positive* requirements that are anchored in law, regulation or professional standards and lack deeper justification. Yet many of these communicative norms are vital for democracy. Prioritising only the two norms for speech that are seen as human rights acknowledges the claims of speakers, but marginalises those of readers, listeners and viewers.

7. Accountability and Anonymity

How then can we support trustworthiness and accountability in a digital age? It seemed until recently that the major digital companies were pushing an agenda of self-regulation that was unlikely to be compatible with accountability for providers and unlikely to secure trustworthiness for others. Many of the large digital companies in the West enjoy concentrated wealth, but have limited physical

[18] Civility may be making a comeback. See recently Teresa Bejan, *Mere Civility: Disagreement and the Limits of Toleration* (Cambridge: Harvard University Press, 2017); Keith Thomas, *In Pursuit of Civility: Manners and Civilization in Early Modern England* (Yale University Press, 2018).

[19] Consider the traditional injunction 'Üb immer Treu und Redlichkeit', from the poem by Ludwig Christoph Heinrich Hölty that Mozart incorporated into the *Magic Flute*. It would hardly do to replace this classical injunction with the injunction 'Üb manchmal Treu und Redlichkeit'! Truth-telling is seen as a *perfect duty*, and not as an optional matter for individual choice. It may be qualified only to satisfy a comparably weighty duty

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

footprints; some are headquartered in jurisdictions with low tax rates, and even more in jurisdictions that are unlikely to bear down on them with anti-trust measures. They appear well-placed to resist attempts to increase their accountability. And the Chinese digital companies have shown all too clearly that digital technologies are ideal for state surveillance and control. A good deal of the evidence pointed to a future dominated either by surveillance capitalism or by surveillance states.[20]

However, as I see it, many of the problems with securing accountability for trustworthy action in a digital age arise not merely from the subordination of digital technologies to corporate and state power, but also from the fact that much that they do is anonymous. Citizens in democracies need to share a common space in which debate and communication takes place, and this is likely to be damaged if anonymous interventions in democratic politics can be undertaken by wealthy individuals (who need not be citizens), or by anonymous corporate and state actors, including hostile state actors.[21] Yet at present digital providers permit all of these to intervene covertly.

Some of the new intermediaries[22] enjoy remarkable combinations of advantages. They control technologies that can deliver content to readers, listeners and viewers without being required to provide those recipients with information about the intermediaries, the originators, the targeting of that content, or information about the source of funds. They may use proprietary algorithms to target selected content on specified audiences without indicating who paid for what and without providing an imprint.

The new intermediaries take little responsibility for the content posted, and are not regulated as publishers. They may (but sometimes do not!) provide anonymity for those whose data they collect, use and market. But they do provide anonymity for those who purchase advertising services that target others. Traditional electoral regulators – for example, in the UK the Electoral Commission – lack powers and capacity to regulate anonymous interventions in elections and referenda by those who may represent no citizens and no political

[20] See Soshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (Profile Books, 2019).

[21] On the latter see Gordon Ramsay and Sam Robertshaw, 'Weaponising News: RT, Sputnik and Targeted Disinformation', (https://www.kcl.ac.uk/policy-institute/research-analysis/weaponising-news, February 2019).

[22] See Srnicek, *op. cit.*

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

parties, and may be orchestrated by hostile state and corporate powers.[23]

As the old intermediaries were bypassed or eliminated, new intermediaries that are neither regulated nor professionalised increased their control. They included coders, data analysts, 'influencers', bloggers, the practitioners of adtech, the customers of the digital providers, and many others who coordinate the provision of platforms and the targeting of persuasive content. Unsurprisingly these new intermediaries can ignore established ethical and epistemic norms and standards, and can disperse activity and limit their own accountability. They may circulate fake, false and flaky content,[24] enabling their customers to exercise covert influence.[25] Today we can look back on the cyber-romantic fantasy that disintermediation would foster democracy with amusement tinged with bitterness. But it is hard to identify remedies.

8. Damage to Communication

It has now become easy, all too easy, for more people to express themselves and to transmit content digitally, but this gain is often not matched by accountability or trustworthiness, since many recipients are unable to judge the claims they encounter. This should not surprise us: digital technologies are excellent for *expressing* and *disseminating content*, but not for *receiving and judging content*, hence not for communication, or for meeting the ethical and epistemic standards that matter for communication and for democracy.

Digital technologies may sometimes extend individual autonomy (at least for consumer choices), but need not support public goods, such as trustworthy journalism, shared cultures or democratic debate and deliberation. All too often, content is originated and organised anonymously, thereby short-changing recipients, who cannot tell who initiated speech acts that are distributed by unknown algorithms, whose inventors and controllers remain hidden from them, and whose specific uses may have been paid for anonymous others.

[23] See the report of the UK Electoral Commission (https://www.electoralcommission.org.uk/__data/assets/pdf_file/0010/244594/Digital-campaigning-improving-transparency-for-voters.pdf). See also DCMS Select Committee Report.

[24] See Onora O'Neill, 'Facebook's trust rankings can't be trusted' (https://www.washingtonpost.com/news/theworldpost/wp/2018/09/25/facebook-2/?noredirect=on&utm_term=.07b9c74834b6).

[25] Cf. Roger McNamee, *Zucked: Waking Up to the Facebook Catastrophe* (Harper Collins, 2019); ; Peter Pomerantsev, *This is Not Propaganda: Adventures in the War against Reality* (Faber and Faber, 2019).

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

So it is hardly surprising that recipients of online content often cannot judge whether the claims they encounter are accurate, honest or reliable, or who or what may be 'promoting' (or inventing!) a given 'message', or suppressing , mocking or deriding other 'messages'. At present recipients of digital content may be unable to tell *who* produced specific claims, *what* they aimed to achieve, *how* they were funded or *what* their underlying aims and agenda were. We face a new version of the predicament that led Socrates to shun writing.

Anonymity is sometimes acceptable, and even necessary. It is needed where repressive regimes or powerful companies seek to prevent investigation or reporting of their action and policies. Investigative journalists and whistle-blowers sometimes need anonymity, not merely to protect their sources, but to protect themselves. But technologies that spread the cloak of anonymity *indiscriminately* for *undeclared* or *covert* purposes can damage and undermine public discourse and thereby also cultural, social and democratic life. Anonymity is not an unconditional good. Freedom of expression, like other human rights, is a *qualified* right that must be configured in ways that do not violate other rights.

9. Regulatory Remedies?

Are there remedies for these problems? Can the familiar disciplines of law and regulation help ensure respect for norms and standards for speech in the digital age? How might this be done?

Regulation of communication sometimes works. For example, in the UK and elsewhere there are numerous regulators of communication, each focused on securing respect for *some* norms and standards in *some* contexts The *Electoral Commissioner* regulates electoral campaigns and expenditure during those campaigns by political parties.[26] The *Advertising Standards Authority* regulates print advertisements and requires an imprint that reveals that certain content is indeed advertising.[27] The *Information Commissioner* regulates compliance with data protection principles.[28] *Ofcom* regulates the television, radio,

[26] For the *Electoral Commission* see https://www.electoralcommission.org.uk/our-work/who-we-are.

[27] For the *Advertising Standards Authority* see https://www.asa.org.uk/

[28] For the data protection principles that *ICO* oversees see https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/. For a recent report see

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press

telecoms and postal sectors and has a statutory duty 'to represent the interests of citizens and consumers by promoting competition and protecting the public from harmful or offensive material'.[29]

However, none of these regulators can regulate speech that is either anonymous or controlled from other jurisdictions. Nor can any of them secure respect for the full range of norms and standards that matter for communication that supports the judgements of trustworthiness needed in public and democratic life. Algorithmic distribution of content may hide originators, paymasters and beneficiaries and may serve undeclared purposes.

The content individuals post deserves the same, but only the same, protection as other individual speech acts. But the amplification, distribution and microtargeting of selected content has other origins, that reflect the business interests, advertising targets, technical standards, and political agendas of those who organise, fund and profit from the dissemination and direction of content. The case for according them anonymity, and in consequence exempting them from accountability, is hardly obvious.

At present it is unclear how much public policy can do to limit the public harms to which digital technologies may lead. Despite moves in the EU and worries in many quarters, the feasibility of legal and regulatory disciplines that will limit damage to culture and democracy remains unclear. It has been widely suggested that what is needed is to bring the tech companies under the same requirements as publishers, and to ensure that they are not treated in law as platforms for others to express themselves, thereby evading responsibility for the targeted content that they distribute. However, this is easier said than done. The leading digital companies combine vast resources with minimal physical infrastructure. The ease with which they can shift jurisdictions and the difficulty they would have in meeting the responsibilities of publishers suggest that we should not assume that change will be easy. But if we want to protect trustworthy institutions and to preserve democracy, change is needed.

ONORA O'NEILL combines writing on political philosophy and ethics with public life. She has been a crossbench member of the House of Lords since 2000, and is an Emeritus Honorary Professor of Philosophy at Cambridge.

Democracy Disrupted: Personal Information and political Influence (ICO, 2018 https://ico.org.uk/media/2259369/democracy-disrupted-110718.pdf).

[29] For *OFCOM* see https://www.ofcom.org.uk/home.

https://doi.org/10.1017/S0031819119000457 Published online by Cambridge University Press