Full Fact response to the Online Harms White Paper

Summary

1. There is a great deal we welcome in this White Paper and we recognise the quality of the work that has gone into it. The goals are important and in many places urgent, and the emphasis the government has placed on protecting free speech and the democratic process is important both in its own right and as an international example.

2. Much of what falls within the White Paper's scope goes beyond Full Fact's charitable mission and expertise. Our mission concerns information quality, misinformation, and disinformation. That takes us into questions of the roles of different actors in acting on online harms and opportunities, but only has limited overlap with other specific harms and opportunities.

3. Full Fact has called for the policy response to misinformation and disinformation to be carried out through open transparent democratic processes. We do not believe that the white paper delivers this. We believe that role of the proposed regulator is far too broad. The parliamentary process exists to protect free speech and civil liberties.

4. We believe that it is important that the powers and duties of internet companies are defined in law. This is clearly a new category of business, and one which has new kinds of powers in new places. It is the normal business of regulation to ensure that powers are scrutinised appropriately and not abused. However, the use of the phrase 'duty of care' is misleading, and the appropriate body to confer new legal obligations on any person is Parliament.

5. There are important differences between tackling disinformation online, with its direct connections to the democratic political process, and other online harms and opportunities. At some points the White Paper does not seem to reflect these as carefully as we feel is necessary. For example, we do not believe that restriction of scope to user-generated content and user interaction is appropriate for tackling the harms from disinformation and misinformation, and we do not believe that the risks associated with interference in free speech have been adequately addressed in the White Paper.

Contents

The legislative structure

Overview: the need for an open democratic transparent processes

There is a great deal we welcome in the White Paper. We would like to thank the officials and parliamentarians involved.

The rise of the internet brings new concentrations of new kinds of power. Citizens in a democracy have a pressing interest in making sure that neither internet companies nor governments misuse these new powers, and in making sure that neither internet companies nor governments nor regulators take decisions unscrutinised that cause problems for others.

The White Paper sets appropriate policy goals and priorities of ensuring the safety of users, protecting freedom of expression, providing clarity for businesses of all sizes, and proportionality. (As noted in section 4, charities run online services too.)

We welcome again the government's explicit commitment to protecting freedom of speech and its record so far of not rushing to legislative action that may have unintended consequences. We recognise that not all governments around the world have taken the same care.

We also believe that the time for action is now. We cannot of course expect that all the decisions taken now will stand the test of time, but inaction is a decision too.

So we urge a realistic approach. Disruptive change in communication technologies takes time to adjust to. However well thought out and careful of civil liberties it is, any legislative framework that results from the White Paper will need to be updated, and principles that are widely agreed now may yet be overturned in future. Both the development of the printing press in the 17th century and the rise of newspapers in the 19th century led to decades of legislative action. The internet is surely as profound a change. The fact that the debate will and must continue is not an argument for doing nothing now. It is an argument for careful action.

Full Fact strongly believes that initiative in this area must come through open transparent democratic processes. In a democracy this means above all public debate conducted by the elected representatives of the people. The parliamentary process is the process we trust to protect free speech, civil liberties, and human rights in the UK. There is no need to invent a new mechanism to do this.

The idea that this process can be outsourced to an appointed regulator is untenable, yet that seems to be the thrust of the White Paper. The so-called 'duty of care' is broad and vague, and the scope of the regulators' suggested powers is broad and threatening. It cannot be adequately managed by a general duty to take a "risk-based and proportionate approach".

Full Fact's biggest concern about the White Paper is therefore the excessively strong and unrealistic expectations of what role the regulator can play. Defining and conferring powers and duties on individuals and organisations is the role of parliament. It is also parliament's role to protect civil liberties and human rights. However, what the government has proposed instead seems essentially

to be that both government and parliament should delegate the whole business to an unspecified regulator.

The proposed delegation of so much discretion to a regulator will obscure responsibility for decisions that may shape society for generations, and which fundamentally and unavoidably are about the trade-offs between different values, freedoms, and rights. It will probably also mean that questions that should be argued out in the democratic process in parliament will instead be argued out in the courts.

The words of Lord Hoffman in the judgement in ex p Simms are a good guide to how to approach where responsibility should lie in this case –

> *"Parliamentary sovereignty means that Parliament can, if it chooses, legislate contrary to fundamental principles of human rights. The Human Rights Act 1998 will not detract from this power. The constraints upon its exercise by Parliament are ultimately political, not legal. But the principle of legality means that Parliament must squarely confront what it is doing and accept the political cost. Fundamental rights cannot be overridden by general or ambiguous words. This is because there is too great a risk that the full implications of their unqualified meaning may have passed unnoticed in the democratic process."*

Parliament should expect to be legislating regularly in this area, which would have several advantages –

1. Clear democratic accountability
2. Parliament builds up its expertise and scrutiny capability
3. The law might not actually go out of date
4. Parliament has authority which no other body can have and primary legislation is less vulnerable to expensive lobbying by litigation

The duty of care

As an ethical statement, of course internet companies have duties to care for their users and others. When their behaviour has suggested they do not recognise this they have at times fallen well short of what we should expect from any reasonable business.

However, 'duty of care' is legal term with a specific meaning. The use of it in the White Paper to describe something different is confusing and misleading. This language should be dropped.

The government should be clear about the reality of what it is doing, which should be proposing new laws for a new world, which is its job.

However, the 'duty of care' seems to be designed to avoid this clarity.

The role of the regulator and the Codes of Practice

The White Paper appears to envisage an extraordinarily broad role for the regulator. It is both a dangerous remit and an impossible task.

The assurance that the regulator will take a risk-based and proportionate approach is an empty assurance when we know that both what counts as a risk worthy of state action, and what counts as a proportionate response, are questions of principle that are, rightly, highly contested political choices.

It is hard to imagine another area where this approach would be considered reasonable. In places it reads as if the government planned to do the equivalent of giving the police the ten commandments and asking them to flesh out the details of criminal law.

We see a great deal of value in the proposed Codes of Practice but think they should fulfil a slightly different role than the White Paper suggests.

The White Paper cites the experience of Corporate Governance Code. We do not accept this analogy. The Corporate Governance Code is essentially a good practice document, and the expectation is that companies (with premium listings) comply or explain. This is a long way away from setting out in clear transparent enforceable ways how to fulfil legal duties, which is what the White Paper says the regulator should be doing.

We think the better analogy is with the ACAS Codes of Practice, which exist to help employers fulfil and demonstrate that they have fulfilled their employment duties as set out in law.

The duties organisations have should be set out principally in primary legislation, as employment duties are. The regulator should set out in its Codes how relevant organisations and services can comply with them. Like the ACAS Codes, a failure on the part of any person to observe any provision of a Code of Practice should not of itself render them liable to any proceedings—but it should be taken into account during any proceedings based on a breach of the duties set out in law.

Disinformation and misinformation

Online harms and disinformation and misinformation

Disinformation and misinformation sit awkwardly in this White Paper. It is right that they are addressed, but it is vital that we are clear about the limits of applying general online harms approaches to this specific area.

On the one hand, online harms are only one aspect of a response to disinformation and misinformation, a set of problems that affects everything from international relations to conversations with your doctor.

On the other hand, not all disinformation or misinformation is, or leads to, harms worthy of government action. Most people's online dating profiles are disinformation by the government's definition. More seriously, although the government is right to include political gain in their definition of disinformation, they are also right to recognise the many risks that come with government action around political speech.

There is only partial overlap between online harms and disinformation so, for example, interventions and remedies that would be uncontroversial in respect of illegal content might be inappropriate for disinformation. In trying to provide a single framework for online harms the White Paper does not always demarcate these boundaries, although we know the government recognises that they exist. That may work in a White Paper but it will not work in legislation, where it will be necessary for example to specify that some power cannot be used in respect of political speech.

Full Fact therefore recommends that the government sets out its thinking and plans on disinformation and misinformation in greater detail, perhaps in a disinformation strategy. We believe this would help the government, other actors in this field, parliament, and the public.

From outside, it appears that government work on disinformation is hindered, or at least not helped, by the low quality of public policy discussion on disinformation and misinformation, exacerbated by the low—and sometimes shockingly low—standards of evidence in areas of this field, along with an evidence base skewed towards what is easiest to research.

We observe gaps in understanding between those looking at disinformation from a national security perspective and those starting from other perspectives, such as media, education and internet, as well as specific topics like public health. This leads to relatively isolated communities of expertise that could, and should, be more effective at working together and learning from one another. We would urge more interdisciplinary work among these fields to identify gaps in evidence and share research methods and best practice.

We see an opportunity for the government to improve the quality of debate and policy making in this area by setting out the landscape as the government sees it, including what we do and do not know. This should be backed by supporting the development of a comprehensive, useful, and accessible evidence base by working with UKRI, learned societies and other research institutions, to develop targeted cross-cutting research programmes.

The role of the regulator

Disinformation and misinformation touch on very many areas of government responsibility, including national security, relationships with other states, counter-extremism, the education system, public health, emergency preparedness, financial stability, consumer protection, and other important public functions. To view this all through the lens of the responsibilities of internet companies would not be credible. In the absence of a wider published strategy on disinformation and misinformation, that is what the government risks appearing to do in this White Paper.

The breadth of concerns in this area, and the need for complex cross-government coordination, raise the question of what role the proposed regulator can credibly play on disinformation and misinformation, particularly given its many other important responsibilities. Arms-length bodies are not usually the best vehicle for cross-government coordination.

From that proposed regulator's point of view, we believe there is a risk that disinformation will quickly become the unloved stepchild among online harms. It raises nuanced and very difficult questions of policy and principle; it is politically highly sensitive; there are many powerful stakeholders who will not be easy to satisfy; there is no clear measure of success; and there are many different facets to the problem, many of which do not have obvious credible interventions waiting to be adopted. It is always going to stay in the 'harms with a less clear definition' box, because it is not one problem and not one harm.

What the regulator can usefully do, which the White Paper touches on, is break down the problems of disinformation and misinformation and identify those areas where proportionate action is needed and justified. For example, it is much easier to get consensus on (some) interventions against false information about public health than some other forms of disinformation.

However, to do this the regulator will need extremely careful attention to public trust, and it should be subject to particular scrutiny. One risk is simply that any appearance of a 'Whitehall knows best' approach is likely to be unpopular. As such, we emphasise that public engagement and not just stakeholder engagement should be one of the key principles of any regulator.

Suggested expectations on internet companies

We welcome this list and think it identifies a good and thoughtful set of areas for expectations which could in some cases be adopted now and in some cases could with further work turn into helpful and appropriate duties for providers of online services.

As a general point, we do not see how it is helpful to pretend that all of the suggested steps listed in the White Paper are simply consequences of a general duty of care.

To Full Fact, the proposed expectations listed in S7.28 of the White Paper fall into four groups, which we discuss in more detail below –

1. Five sensitive areas
2. Transparency measures which we welcome
3. Fact checking
4. News quality

Five sensitive areas

Each of these identifies an important area. As the White Paper provides for, further discussion is needed to develop clear duties which, in our view, should then be set out in law.

- The steps companies should take in their terms of service to make clear what constitutes disinformation, the expectations they have of users, and the penalties for violating those terms of service.
- Steps that companies should take in relation to users who deliberately misrepresent their identity to spread and strengthen disinformation.
- Improving the transparency of political advertising, helping meet any requirements in electoral law.
- Companies will be required to ensure that algorithms selecting content do not skew towards extreme and unreliable material in the pursuit of sustained user engagement.
- Promoting diverse news content, countering the 'echo chamber' in which people are only exposed to information which reinforces their existing views.

On transparency of political advertising, we have set out before the need for full transparency of content, targeting, reach, and spend in machine readable formats in real time. The initiatives taken by the major internet companies are welcome but not enough.

On the final point, there is of course value in offering people diverse news content, but we are wary of the government's attempt to counter 'echo chambers', especially given the lack of robust evidence on the topic. Platforms have made numerous attempts to limit harmful information, and promote trustworthy sources, and indeed the latter of these is addressed elsewhere in the suggested codes of practice.

Transparency measures which we welcome

We are cautiously optimistic about measures that focus on transparency, through better and clearer reporting processes and increased expectations around monitoring and evaluation.

- Ensuring that it is clear to users when they are dealing with automated accounts, and that automated dissemination of content is not abused.
- Processes for publishing data that will enable the public to assess the overall effectiveness of the actions companies are taking, and for supporting research into the nature of online disinformation activity.
- Reporting processes which companies should put in place to ensure that users can easily flag content that they suspect or know to be false, and which enable users to understand what actions have been taken and why.
- Steps that services should take to monitor and evaluate the effectiveness of their processes for tackling disinformation and adapt processes accordingly.

Mandating the use of fact checking services

- Making content which has been disputed by reputable fact-checking services less visible to users.
- Using fact-checking services, particularly during election periods.

We are grateful for the government's recognition of the value of fact checking and we are keen to do more of this kind of work where it has a clear public benefit. However, the government should always be cautious about creating markets by regulatory fiat.

In implementing a duty in this area it would be valuable to consider the steps a service can take to affect the reach of the fact checking among its users, and how this and other outcomes can be independently scrutinised.

The government should also recognise that the existence of high-quality independent fact checking organisations is not inevitable. Funding work like Full Fact's is extremely difficult. It would not be desirable for any company to be mandated to work with organisations that do not exist.

One possible unintended consequence of these expected steps is to promote new entrants to fact checking with lower standards and a pure profit motive rather than a public benefit mission. This could undercut the work of Full Fact as a charity and the work of media outlets doing high-quality fact checking. Ultimately, this risks reducing the quality of fact checking done online, thus damaging the public's trust in the process of fact checking.

For background, the Code of Practice of the International Fact Checking Network sets out a minimal baseline standard of transparency for fact checkers, and Full Fact is a signatory. Full Fact is also the working with Facebook's Third Party Fact Checking initiative (we go into this in more detail in the answers to questions section below) and we will be publishing our first report on its operation in the coming weeks.

News Quality Obligation

- Promoting authoritative news sources

We believe it is important that people have access to trustworthy news and we know more than most about the difficult of making those distinctions. That's not just the vexed question of deciding that some news outlets can be considered reliable and some not; it also extends to questions like ensuring media outlets promoted represent a diverse range of journalism that serves a diverse range of audiences (local, demographic, etc.). There is a reason the BBC doesn't just offer Radio 4.

We have not yet seen an approach to identifying authoritative news sources that we are confident goes beyond a complicated way of creating a 'whitelist' of news sources somebody approves of. This includes efforts to determine a set of indicators and signals that can be used to assess trustworthiness automatically. If creating a whitelist is what is being done, it would be better to be honest about it, particularly in a country which already has public service broadcasters with mandated standards of accuracy and due impartiality.

All of these projects tend to be constrained by the fact that, practically speaking, they have to end up with all or almost all existing major media outlets being considered trusted. It is easy to see how this can end up baking-in assumptions that ultimately act as a barrier to entry to new players at a time when we have more new players than ever.

The White Paper specifically refers to NewsGuard. Trying to create a business by persuading governments to legislate the market into existence is a well-worn tactic. Again, the government

should always be cautious about creating markets by regulatory fiat and we do not believe that it should in this case. We are not satisfied either that there is an adequate evidence base for this intervention, or that the model of domain-level judgements is credible, or that the content provided by NewsGuard is sufficiently high quality to justify its adoption.

Difficult judgement calls

In Section 7.31, the government states: “Importantly, the code of practice that addresses disinformation will ensure the focus is on protecting users from harm, not judging what is true or not. There will be difficult judgement calls associated with this. The government and the future regulator will engage extensively with civil society, industry and other groups to ensure action is as effective as possible, and does not detract from freedom of speech online.”

This suggests that the government has not engaged with what this would mean in practice, at least in respect of disinformation and misinformation.

These “difficult judgement calls” are essentially about freedom of speech, freedom of association, and the freedom to impart and receive information, and their limits. The suggestion that the right balance should be determined by a government-appointed regulator working with “civil society, industry and other groups” makes for uncomfortable reading. This is not a conversation for nerds, wonks and businesspeople in closed rooms.

These trade-offs need to be made through an open transparent democratic process, in parliament, and not a technocratic process in a regulator’s office.

Even if the government rejects this view and parliament fails to insist on it, we would at the very least expect an explicit commitment to public engagement and a plan for how it will work.

Technology

We welcome the government’s recognition that technology can play a crucial role in keeping users safe online. We note that Mark Zuckerberg has similarly said: “we’re going to shift increasingly to a method where more of this content is flagged up front by A.I. tools that we develop.”

Full Fact are world leaders in using AI for fact checking and have recently won the Google AI for Social Good Impact Challenge along with our partners AfricaCheck, Chequeado, and the Open Data Institute.

Disinformation and misinformation are not one problem. When you break them down into smaller problems, some are solvable, and some are solvable using technology. Full Fact is working on automated fact checking to identify those parts and to use technology to scale, target, and dramatically increase the effectiveness of our work, and we are already seeing the benefits.

However, at the moment, we would be deeply sceptical about any technology claiming to be able to distinguish trustworthy and untrustworthy information at scale in such a way that it could be used to promote or demote arbitrary content being shared online. Even if a technology appeared to deliver useful results, we would expect to find damaging unintended consequences when we scrutinised how it worked.

Technology innovation in this area needs a public benefit focus, scrutiny and accountability. Care needs to be taken about who is disadvantaged by the unintended consequences of the tools. There are not enough actors in this space who have this kind of approach.

There are two areas Full Fact believes the government could invest in to encourage a step change in the quality of work in this area –

1. Review and support the development of better Natural Language Processing data and libraries for more languages, perhaps through the international development budget. Existing tools provide good support for some Western European languages, less support for Asian languages, and generally favour the languages of richer countries. This means that we are some way from being able to provide global technological solutions to what is a global problem. These limitations have hindered Full Fact's and our international peers' work in this area and are a key barrier to the wider rollout of tools we already have.

2. Support fact checkers and domain experts to provide independent assessments and benchmarks of proposed technologies by probing their strengths and weaknesses against carefully-designed test inputs. This function would be analogous to the role that Euro NCAP, the keeper of the crash test dummies, plays for the car industry.

We would be glad to discuss this further if there is any appetite for future work.

Answers to questions

Question 1: This government has committed to annual transparency reporting. Beyond the measures set out in this White Paper, should the government do more to build a culture of transparency, trust and accountability across industry and, if so, what?

The requirement for annual transparency reporting is welcome, and the measures outlined are a sensible starting point. We particularly welcome commitments to ask for reporting on safeguards to uphold and protect fundamental rights, and details of investment in support for user education and work with civil society.

But we would argue that, in order to fully understand and evaluate platforms' actions, more information should be collected in three areas.

First, there is an urgent need for transparency about political advertising . We have consistently called for a political ad database that is provided in real-time, year-round and in machine readable format. This should contain full information on an ad's content, the group it is targeted at, its reach and the amount spent on it. This should be publicly accessible rather than controlled by a private company.

Second, platforms need to provide more data to academics and factcheckers to allow for independent evaluation and greater transparency of practices. Although there are some voluntary efforts, as the White Paper states, these do not go far enough or fast enough and are not applied consistently across platforms.

Other efforts open up only some data to select groups of researchers. Such piecemeal access risks creating a two-tier system of understanding and reduces the ability for wider conversation and analysis. The government should require that annual reports include details on which researchers have been granted access to what data, how it has been used and where research based on that has been published. Ideally, the regulator should encourage the platforms to require that this research be published in open-access journals.

Full Fact is the UK's partner in Facebook's Third Party Fact-checking scheme, and as part of this we have committed to regular reporting on our work. The first report will be released shortly. It is also vital that the database of factchecked articles that Facebook gathers from the scheme, and any tools developed using this, are given external scrutiny to ensure they are ethical, fair and responsible. The government could call for similar commitments from Facebook.

Third, platforms should be wary of algorithmic approaches to identifying misinformation. While artificial intelligence technology can help humans spot patterns of behaviour or patterns in content, in the field of disinformation and misinformation it remains imprecise and should not be relied upon to do the job of human moderators. Planned reporting on the use of technological tools should probe the platforms' balance of human and algorithmic moderation.

We also believe that transparency and accountability should go both ways, and while the government rightly plans to seek more openness from the platforms, it should be willing to open up some of its own data. For instance, it would be beneficial for researchers to be able to request

information on algorithms developed by companies as part of government-backed projects to tackle disinformation or misinformation.

Finally, any discussion about how to improve trust and accountability in the system should consider the importance of transparency in how decisions about regulation are made. We welcome this open consultation, and would ask the government to set out how it plans to involve the public in the future work of the regulator.

Question 2: Should designated bodies be able to bring 'super complaints' to the regulator in specific and clearly evidenced circumstances?

Yes.

Question 2a: If your answer to question 2 is 'yes', in what circumstances should this happen?

Complaints are capable of being used as a lobbying mechanism and any procedure for designating bodies capable of making super-complaints should reflect this risk.

The White Paper implies that super-complaints should exist "to defend the needs of users". We think there would also be circumstances where they should be available to defend the needs of non-users. It should be a public interest test, not just the equivalent of a consumer protection one.

The criteria under The Police Super-complaints (Criteria for the Making and Revocation of Designations) Regulations 2018 are largely appropriate for this context, as is the procedure for complaints under the Police Super-complaints (Designation and Procedure) Regulations 2018.

The practical constraints under this procedure are discretion about designation and discretion about how any specific is handled.

Question 3: What, if any, other measures should the government consider for users who wish to raise concerns about specific pieces of harmful content or activity, and/or breaches of the duty of care?

In the case of disinformation and misinformation, this is a difficult question. We share the government's view that "applying 'publisher' levels of liability to companies would not be proportionate", and its reasons. It follows that we do not expect internet companies to routinely review individual content just because someone complains about their accuracy.

We note that some of the most prominent online services do provide a facility for users to flag content they believe is false. We don't know how often this is used or what happens to that information. For all we know, in some cases it might be the equivalent of the button in the lift that does nothing except make the user feel better. People might at least be entitled to an explanation and we welcome the suggested expectation of specifying: "Reporting processes which companies should put in place to ensure that users can easily flag content that they suspect or know to be false, and which enable users to understand what actions have been taken and why."

In fact, the responsibility for the accuracy of what is posted online usually belongs to the poster. The platform can make some choices that nudge posters' behaviour in one direction or another. They can make a facility for sharing corrections available; they can recognise or even reward publishers

who show a commitment to accuracy including correcting mistakes; they can make sure it is possible to link to sources and reward that behaviour.

They can go further and work with third-party fact checkers and share our content, as discussed above. If they choose to do so they should also ensure that the results are shared to those who have shared and seen the content.

Question 4: What role should Parliament play in scrutinising the work of the regulator, including the development of codes of practice?

We have answered this more extensively above. If the government accepts the case for the Code of Practice to play a more limited role than the White Paper suggests, then the procedure adopted for ACAS Codes would be sufficient.

If the Code of Practice has the role the White Paper seems to envisage at the moment it is not clear that any level of parliamentary scrutiny short of primary legislation would be adequate. If a middle ground is reached, then the super-affirmative procedure for delegated legislation under the Legislative and Regulatory Reform Act 2006 might provide a model.

Question 5: Are proposals for the online platforms and services in scope of the regulatory framework a suitable basis for an effective and proportionate approach?

We are pleased that the government has recognised that this framework will have to apply to a broad range of companies across many sectors, and that there is an understanding that it has to be aimed at more than just the main online platforms.

The restriction to services hosting user-generated content or user interactions does not work for tackling the harms from disinformation and misinformation online. There is no basis for believing that these are specific to user-generated content.

Question 6 In developing a definition for private communications, what criteria should be considered?

Question 7: Which channels or forums that can be considered private should be in scope of the regulatory framework?

Question 7a: What specific requirements might be appropriate to apply to private channels and forums in order to tackle online harms?

The terms ‘private communications’ and ‘private channels’ suggests that there is a bright line to be drawn between what is private and what is not. We do not think there is. The government acknowledges this complexity in the White Paper.

The difficulty is with the government’s statement that: “users should be protected from harmful content or behaviour wherever it occurs online, and criminals should not be able to exploit the online space to conduct illegal activity.”

We cannot see how this well-intentioned ambition can be achieved outside of a totalitarian system. No government has taken the view that harmful speech should not be allowed offline: the criminal

law stops well short of that. The criminal law is also not designed to be perfectly and automatically enforced offline, and if it was there would serious unintended consequences.

Again we stress that ordinary people getting things wrong online is not a harm that requires government action.

Now that it is technically possible for government to monitor private communication at scale there will inevitably be policy reasons to do so. There are also policy reasons not to do so in an open society. It is a political judgement where the balance lies, and the choice should be made through an open democratic transparent process by parliament.

Full Fact is watching international examples of harmful disinformation and misinformation spreading via non-public channels with concern. We do not know how far this affects the UK yet, but we assume it is at least part of the future. One response to that might be ever-greater interference in people's private communications. We expect that the better response will be more investment in communication reliable and trustworthy information. An open society should use debate, not control, to respond to disinformation.

The challenge for anyone who believes that direct intervention in non-public conversations is an appropriate response to disinformation and misinformation is to show that other options are not sufficient. Given that we can measure public knowledge, beliefs, and attitudes - and given that we can communicate with the public - why should it be necessary to seek to control the conversation?

That said, we of course recognise that there are times when interference with free speech can be proportionate, and so does the European Convention on Human Rights, and so does the US Supreme Court. History shows that argument is seductive. Even the strongly-worded "clear and present danger" test for interfering with free speech became, in the words of the US Supreme Court "manipulated to crush" people's freedoms, by or with the support of the courts themselves.

So, if the government proposes to intervene against disinformation in non-public channels it should set out in legislation a high threshold for doing so. The more that authority is delegated, the higher that bar should be. This risk would matter a little less if the government were not proposing to treat such sensitive questions as administrative matters to be outsourced to an arms-length body.

Our suspicion at the moment is that the government will be tempted to try to undertake censorship-by-proxy. The government will have legitimate, important, and life-saving policy concerns—for example about vaccine take up. They will lean on the regulator, who will lean on the internet companies, who will amend their terms and conditions and take action. We understand the real harms caused by false beliefs about vaccines and other topics, and the good intentions that would drive this process. The government needs to be clear about the cost of such a profound intervention in public debate being made other than through an open democratic transparent process. We do not believe that price is worth paying.

Question 8: What further steps could be taken to ensure the regulator will act in a targeted and proportionate manner?

Above all, that more of the powers and duties are set out by parliament in primary legislation.

Once established, a vital part of ensuring the regulator acts in a targeted and proportionate manner will be to make clear distinction between the different forms of harm, and sub-categories of harm, acknowledging that some cases will require tougher actions from online platforms and the government than others.

For instance, there is a distinction between moderating abusive content and moderating false or misleading content – the latter is where Full Fact can offer expertise. There is a proportion of inaccurate material than can be cleaned up simply, like spam. But the more vigorous the efforts of online platforms to counter misinformation, the greater the risks to freedom of expression. The focus on fairness as well as effectiveness in the question is vital.

It will also be crucial that the Codes are written in such a way that they take into account how much we do, or don't, know about the scale and spread of misinformation and the actual harms it causes. It is welcome, therefore, that the White Paper suggests the regulator has greater collaboration with UKRI, and Full Fact would urge that this extends beyond research council funded academics to other researchers and civil society groups working in the field. We are already leading an international research project in this area and we believe there are great opportunities in a more coordinated approach.

Question 9: What, if any, advice or support could the regulator provide to businesses, particularly start-ups and SMEs, comply with the regulatory framework?

We've suggested following the highly-successful ACAS model with codes and templates that any organisation can freely adopt.

Question 10: Should an online harms regulator be: (i) a new public body, or (ii) an existing public body?

Question 10a: If your answer to question 10 is (ii), which body or bodies should it be?

We do not have a strong opinion, and the government has the experts on the machinery of government, but the internet touches every area of life and the idea of a monolithic regulator for online harms risks trying to create a regulator for life in general. Not only would creating such a regulator be a logistical challenge, it would be expensive and time-consuming.

There is certainly a role for existing regulators that could not easily be subsumed—in our field, that includes at least the Electoral Commission, the ICO, and Ofcom. The government must also recognise that there is a balance to be struck between the existing roles these expert bodies play, and any new ones that are handed to them as a result of this legislation.

If regulators are presented with extra work and limited funding, some areas may be de-prioritised. The risk is that these will be the areas that the government of the day feels less public pressure about, but which are no less vital to the proper functioning and transparency of the UK's systems.

It may be that there is a role for a complementary or coordinating body here, but it is for the government to demonstrate how it will ensure regulators are able to manage their competing interests.

The Electoral Commission at least needs greater funding and access to digital skills to allow it to successfully protect the integrity of our democracy. The Electoral Commission does not have the resources, or therefore the digital skills, to fulfil its duty to review political advertising in broadcast and electronic media, in the way that is needed given the scale, pace, and importance of changes in political campaigning.

Question 11: A new or existing regulator is intended to be cost neutral: on what basis should any funding contributions from industry be determined?

The government may wish to bear in mind that one of the world's top ten websites is run by a charity (Wikipedia, run by the Wikimedia Foundation) and ask whether it intends to add to their costs.

In general, we expect online services to continue to change unpredictably and we suggest that it should be a priority to minimise barriers to entry for new services.

Question 12: Should the regulator be empowered to i) disrupt business activities, or ii) undertake ISP blocking, or iii) implement a regime for senior management liability? What, if any, further powers should be available to the regulator?

This question is a good example of the danger of blurring general questions about appropriate regulation on internet companies with the specific area of disinformation. Whatever the merits of ISP blocking in other context, ISP blocking of political speech to reduce the harms from disinformation is not a proposal we would expect to see in an open society. The government should be much clearer about these boundaries.

Senior management liability is in principle a good idea in some circumstances. False information can ruin and cost lives. Individual liability exists for Health and Safety Offences and in principle therefore it could also apply where there is a similar level of responsibility and harm.

Question 13: Should the regulator have the power to require a company based outside the UK and EEA to appoint a nominated representative in the UK or EEA in certain circumstances?

Yes.

Question 14: In addition to judicial review should there be a statutory mechanism for companies to appeal against a decision of the regulator, as exists in relation to Ofcom under sections 192-196 of the Communications Act 2003?

Question 14a: If your answer to question 14 is 'yes', in what circumstances should companies be able to use this statutory mechanism?

Question 14b: If your answer to question 14 is 'yes', should the appeal be decided on the basis of the principles that would be applied on an application for judicial review or on the merits of the case?

We have argued that the duties applied as a result of the White Paper should principally be set out in primary legislation. It follows that many of them will fall to be enforced by the civil and criminal

courts as appropriate. The regulator might then in some circumstances by the body bringing actions to enforce these duties.

When it comes to decisions made at the regulator's discretion, this depends on the ambit of the regulator's role. Judicial review is unlikely to be daunting for major internet companies, but if the regulator affects a wide range of organisations, judicial review is beyond the resources of many. The other end of the spectrum of risks is of course that regulated entities with deep pockets simply use any appeal mechanism (or the threat of it) to frustrate the work of the regulator.

If judicial review is considered in this case, the government should consider whether there should be costs rules to avoid parties being deterred from bringing legitimate cases by the risk of incurring excessive costs.

The experience of the Charity Tribunal may be relevant. As we understand it, the original thinking was that the Tribunal would act as a light-touch and accessible second look at decisions made by the Charity Commission, where legal representation would not be essential, and that it would become a venue for charity law to continue to evolve. This has largely not happened. Charity law is quirky but relatively well understood. Given this experience, it is hard to imagine who would have to be members of a Tribunal for it to be effective in grappling with the range of issues the regulator is proposed to cover.

Question 15: What are the greatest opportunities and barriers for (i) innovation and (ii) adoption of safety technologies by UK organisations, and what role should government play in addressing these?

The government should ensure that any efforts to develop technologies that tackle online harms are focused on specific problems. Broadly speaking, the more specific the problem, the more likely it is that algorithmic approaches will be accurate, and vice versa. Technological solutions to very broad problems are therefore often not realistic or desirable. For example, we believe that trying to develop tech that baldly classifies content as 'true' or 'false' – known as truth labelling – not only misunderstands the capabilities of the technology, but also the nature of the world we live in.

Full Fact's approach has been to identify solvable problems in factchecking and develop technology to solve those specific issues. Within the global factchecking community, there is a desire for automated tools that can help them detect and check claims, and Full Fact is leading the field in developing these. Widespread adoption of such tools will help ensure that factchecking can be done faster and more consistently, thus providing the public with access to better information on the facts behind mis- or disinformation.

Question 16: What, if any, are the most significant areas in which organisations need practical guidance to build products that are safe by design?

In the ACAS-inspired model we've suggested, it would be the regulator's task to find this out and to provide Codes of Practice that serve these needs.

Question 17: Should the government be doing more to help people manage their own and their children's online safety and, if so, what?

We should be wary of putting too much of the onus on maintaining community standards and responding to the challenges of the internet on users. Our focus should be on trying to help users to make informed decisions, and making those decisions as easy as possible, rather than putting the responsibility for judging content on users.

It will continue to get harder for users to make informed choices about what content to trust—it is now easy to use artificial intelligence to combine and superimpose existing images and videos to create fake videos of famous people. It is difficult even for technical experts to distinguish between the real video and the manipulated video, so there is little hope for the average user.

At Full Fact, as well as publishing factchecks, which set out the evidence on widely debated claims—or state when there is a lack of evidence or data in a certain area—we also have a toolkit to help users challenge claims themselves. This offers step-by-step guides to help them question claims they are seeing, for instance on recognising poorly-created surveys or false images and videos online.

If the government wants to help people manage their safety online, it should consider how to encourage them to question, challenge and debate the information they see, in addition to requiring that platforms have more straightforward approaches for reporting harmful content. In particular, in our work in this area we have started to stress that people who make false news try to manipulate your feelings, and encourage people to consider how what they see makes them feel.

We welcome the planned online media literacy strategy.

Question 18: What, if any, role should the regulator have in relation to education and awareness activity?

This is a specialised skill, and the Government Communications Service and its predecessors have at times excelled in it. Given that Ofcom has an existing duty on media literacy, we recommend the government build on this rather than giving a potentially overlapping duty to any other body. (Full Fact is part of Ofcom's new Making Sense of Media Advisory group.)

However, we believe this work needs to be expanded and made more central to Ofcom's work. For now it is still treated much as if we still lived in a world of four or five TV channels overseen by Ofcom and about ten national newspapers. Media literacy (or news, digital, information literacy etc.) is more complex now and also more needed than ever.

One strength of Ofcom's approach that should be preserved is the emphasis on research and evidence. Media literacy is now a fast moving target and can only be successfully tackled with a useful evidence base.

The government, perhaps through UKRI, should invest in providing rigorous research and evaluation of all projects in this area, including those carried out by third parties from charities to the internet companies (whose cooperation should be mandated by law if necessary), and in sharing the results of those evaluations. This field needs something like the Justice Data Lab and the Education Endowment Foundation's Evidence Summary Toolkits.