The International Journal of Human Rights

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/fjhr20

The right to freedom of thought: an interdisciplinary analysis of the UN special rapporteur's report on freedom of thought

Patrick O'Callaghan, Olga Cronin, Brendan D. Kelly, Bethany Shiner, Joel Walmsley & Simon McCarthy-Jones

To cite this article: Patrick O'Callaghan, Olga Cronin, Brendan D. Kelly, Bethany Shiner, Joel Walmsley & Simon McCarthy-Jones (2024) The right to freedom of thought: an interdisciplinary analysis of the UN special rapporteur's report on freedom of thought, The International Journal of Human Rights, 28:1, 1-23, DOI: 10.1080/13642987.2023.2227100

To link to this article: https://doi.org/10.1080/13642987.2023.2227100

Published online: 29 Jun 2023.

THE INTERNATIONAL JOURNAL OF HUMAN RIGHTS
2024, VOL. 28, NO. 1, 1–23
https://doi.org/10.1080/13642987.2023.2227100

Routledge
Taylor & Francis Group

The right to freedom of thought: an interdisciplinary analysis of the UN special rapporteur's report on freedom of thought

Patrick O'Callaghan[a], Olga Cronin[b], Brendan D. Kelly[c], Bethany Shiner[d], Joel Walmsley[e] and Simon McCarthy-Jones[c]

[a]School of Law, University College Cork, Cork, Ireland; [b]Irish Council for Civil Liberties, Dublin, Ireland; [c]Department of Psychiatry, Trinity College Dublin, Dublin, Ireland; [d]School of Law, Middlesex University, London, UK; [e]Department of Philosophy, University College Cork, Cork, Ireland

ABSTRACT

In 2021, the United Nations Special Rapporteur on Freedom of Religion or Belief presented his 'Report on Freedom of Thought' to the United Nations General Assembly. This was the first substantive consideration of the right to freedom of thought at the United Nations level since the right was recognised in 1948. This paper provides interdisciplinary reflections on this report to support ongoing discussions on the appropriate content and scope of this fundamental human right. We begin by addressing reasons for the historical neglect of this right, namely the right being viewed as more symbolic than practical and relevant interests being perceived as already protected by other rights. Next, given there is no consensus on what the right protects, or how it protects, we consider its potential attributes. We then consider potential violations of this right, turning to its application to mental health. Finally, we consider the Special Rapporteur's recommendations, discussing how some may be realised through human rights-centered regulation in the form of the European Union's new Digital Services Act. In this context, we also briefly consider relevant aspects of the EU Commission's proposal for an AI Act. We conclude by outlining pressing challenges facing the development of this right.

ARTICLE HISTORY
Received 8 February 2023
Accepted 9 June 2023

KEYWORDS
Freedom; thinking; thought; human rights; moral panic; philosophy

1. Introduction

In October 2021, the then United Nations Special Rapporteur on Freedom of Religion or Belief, Ahmed Shaheed, submitted to the members of the General Assembly a 'Report on Freedom of Thought' (henceforth 'The Report').[1] This represented the first substantive consideration of the right to freedom of thought at the level of the United Nations since the right was initially recognised as part of the Universal Declaration of Human Rights (UDHR) in 1948. Thus, whilst The Report did not 'conclusively resolve' ongoing controversies about what constitutes thought or freedom of thought, it did undertake 'the first attempt to comprehensively articulate the right's content and scope in the United Nations system.'[2]

CONTACT Simon McCarthy-Jones mccartsi@tcd.ie

The Report was the result of a consultation exercise that included roundtables and meetings with key stakeholders from across the globe, including legal experts, psychologists, neuroscientists, policymakers, the media, digital technology companies, and both intergovernmental and international organisations. A call for submissions also yielded materials from a range of sources, including civil society entities, individuals, States, and human rights bodies.

The range of expertise consulted by The Report supports the contention that an interdisciplinary approach is required to reach a satisfactory conception of the right to freedom of thought. Such efforts are currently underway.[3] To contribute to this process, this paper will offer a series of interdisciplinary reflections upon The Report, which we hope will be useful to the continuing conversation in this area. These reflections were created by assembling a multi-disciplinary team comprising both academics and practitioners (including those working in the non-governmental sector) working in the areas of law, philosophy, psychology, and psychiatry. The team reflected on the following four key elements of The Report, which, at the same time, provide the structure for our analysis in this paper.

First, the Special Rapporteur finds that the right to freedom of thought has received 'scant attention in jurisprudence, legislation and scholarship, international and otherwise.'[4] This prompts the immediate questions: if the right is so important, why has it received such scant attention? Has the right not been perceived as being threatened to date? Do other rights such as the right to freedom of expression or privacy safeguard the interests at stake? Are there other explanations?

Second, the Special Rapporteur mapped four possible attributes of the right based on international human rights jurisprudence and commentary, namely, not being forced to reveal one's thoughts; no punishment and/or sanctions for one's thoughts; no impermissible alteration of one's thoughts; and States fostering an enabling environment for freedom of thought.[5] In this paper, we seek to assess whether this framework is convincing and sufficiently comprehensive. Can freedom of thought be conceptualised in other ways? Or perhaps we do not even have a sufficiently rigorous shared understanding of what we mean by 'thought' in the first place?

Third, the Special Rapporteur examines potential violations of freedom of thought in seven 'diverse fields': (a) torture or cruel, inhuman or degrading treatment or punishment; (b) surveillance; (c) coercive proselytism; (d) anti-conversion and anti-blasphemy efforts; (e) intellectual freedom and education; (f) existing and emerging technologies; (g) mental health; (h) conversion practices.[6] So understood, the right to freedom of thought has a potentially wide range of application, each field raising distinct issues of concern and controversy. A comprehensive analysis of the Special Rapporteur's findings in each of these areas would not be possible within the confines of this paper. For this reason, we direct our comments towards two of the fields: mental health and existing and emerging technologies. We use the former as a way to interrogate the Special Rapporteur's findings as well as their implications. As for the latter, at various points throughout the paper, and especially in the context of the Special Rapporteur's recommendations, we engage with questions about the potential use of the right to freedom of thought in the field of existing and emerging technologies.

Finally, the Special Rapporteur makes several recommendations to multilateral, State and non-State actors on how to 'respect, protect and fulfil freedom of thought'.[7] At the

heart of the Special Rapporteur's recommendations is what appears to be a plea for a 'human rights-centered design, deliberation, and oversight' of technology.[8] In the final section of our paper, we offer some reflections on how such a human rights-centered approach might be realised in the context of two recent legislative developments in the European Union: the Digital Services Act and the proposal for an AI Act. We will begin our analysis with the first issue: why has the right to freedom of thought been historically neglected?

2. Neglect of the right

As we see it, there are two (potentially overlapping) explanations as to why the right to freedom of thought has received 'scant attention'. First, it may be that the right is considered to hold more symbolic importance than practical value. Second, it may be that interests that fall under its protective remit are already adequately protected by other rights, such as the right to freedom of expression or the right to privacy.

2.1. More symbolic than practical?

Commentators often describe the right to freedom of thought in lofty, exalted terms, seeking to emphasise that it provides a foundation for a range of other rights. Indeed, the drafters of the UDHR claimed that the right to freedom of thought was 'a sacred and inviolable right [and] the basis and the origin of all other rights'.[9] Before this, in the US Supreme Court case of *Palko v Connecticut* (1937), Justice Cardozo had identified freedom of thought and speech as 'the matrix, the indispensable condition, of nearly every other form of freedom.'[10] The foundational role of the right to freedom of thought is emphasised in the opening lines of The Report which refers to 'the essentiality of "freedom of thought" for the dignity, agency and existence of the human being.'[11]

Yet, despite the prominence given to the right in statements of this sort, freedom of thought, understood as distinct from other related rights, rarely features in litigation or as a dedicated subject of study in legal scholarship. This is likely because, as the Special Rapporteur observes, 'the absolute nature of freedom of thought – coupled with what some argue is a narrow scope of protection – has made it difficult to envisage just how and when this right may be violated, thereby undermining its practical application.'[12] We may also add that if thought is viewed as an activity performed in a private, unobservable *forum internum*, there are significant epistemic barriers to knowing when the right to freedom of thought has been violated. Understood in this way, some might regard the right as merely being of symbolic importance rather than of any practical value.

However, the view that the right to freedom of thought has no utility is becoming quickly outdated as new scholarship highlights the ways in which the right might be engaged or at stake.[13] Scholars have linked the right to freedom of *unmanifested* thought to neuroscience,[14] non-consensual therapies,[15] data collection, including in the context of mental health websites, and manipulated (political) decision-making,[16] to demonstrate that the view of unexpressed thoughts as 'intangible'[17] is 'unsustainable'.[18] To take one such practice, The Report identifies existing and emerging technologies as a potential concern for freedom of thought including predictive technologies and

micro-targeting, both based on data gathering, and neurotechnology. The Report refers to a study which used neuroimaging to infer suicidal thoughts with 91% accuracy and studies which make it possible for memories to be modified, removed, or recovered.[19] The Report is concerned that neuroscience can be used to 'sanction inferred thoughts' and even 'modify or manipulate thoughts inside the brain'.[20]

2.2. The interests in question are adequately protected by other rights

Another explanation for the scant attention that freedom of thought has historically received is the perception that other rights already safeguard the interests in question, even if they do so indirectly. Besides the connected rights of *conscience, religion,* and *belief* we may also consider the intimate associations between the right to freedom of thought and the rights to *freedom of expression* and *privacy*. It is also worth mentioning the *prohibition against torture or cruel and degrading treatment or punishment* and the *right to mental integrity* where there is also potential for some overlap with the right to freedom of thought.

In legal systems founded on the rule of law and human rights, the right to freedom of expression tends to be a long-established right with a significant body of jurisprudence and scholarship that has developed around it.[21] The law conceives of our thoughts as being manifested in our speech and/or behaviour, pointing to an important connection between freedom of expression and freedom of thought. But the picture is more complex since the interaction between thought and speech must be understood in its relational contexts. While '[t]hought and expression are conceptually and practically distinct', as the Special Rapporteur points out, 'they engage in a perpetual feedback loop in which expression is a vehicle for exchanging and developing thoughts, and thoughts feed expression'.[22] This more complex picture is also recognised in the legal doctrine since the right to freedom of expression is generally understood as encompassing both the freedom to receive *and* to impart information and ideas.[23] Against this background, it perhaps comes as no surprise that commentators might view the right to freedom of expression as a means of safeguarding basic interests in freedom of thought, at least in some situations.

Before proceeding, two further points can be made about the relationship between the distinct rights to freedom of thought and freedom of expression. First, we should be aware that safeguarding freedom of thought could result in further permissible limitations being created on expression.[24] A crucial question in relation to freedom of thought is where we draw the boundary between permissible influence on others' thinking and what The Report refers to as 'impermissible alteration' of thoughts (a central question, as highlighted by Bublitz).[25] In drawing such a boundary, certain speech acts may be deemed to be violating other people's freedom of thought. At present, in principle at least, if a person's right to be *free from* impermissible influence on their thoughts clashes with another person's right to be *free to* make certain utterances, then the absolute status of the right to freedom of thought means it should always win out in such a conflict. Enforcing an absolute right to freedom of thought may hence chill expression. This causes a further conceptual headache given that, as noted above, free expression is also needed for free thought. In short, although The Report notes that 'infringements on the right [to freedom of thought] could have a chilling effect upon expression',[26] we also

need to consider the potential for the right to freedom of thought itself to have a chilling effect on expression.

Another related issue is that the legal distinction between thought and speech does not map neatly onto psychological conceptions of these phenomena. From a psychological perspective, some speech can be seen, not as manifestations of thought, but as constitutive of thought. In such situations, speech does not express thought but creates it in the moment. Indeed, the Russian psychologist Vygotsky argued that thinking begins as a social act, which is then internalised.[27] In this view, thought begins as an out-loud dialogic exchange between a child and an adult, before going underground as our inner speech.[28] In this sense, thought returns to its origins when we think aloud together as adults. This means, in theory, that what may have traditionally been deemed a permissible limitation on expression could now be seen to be an impermissible limitation on thought. In this sense, the legal and psychological conceptions of thought are notably different. We will discuss this issue of how we should conceive of thought further below.

There is also a close relationship between the right to freedom of thought and the right to privacy (see Article 12 UDHR, Article 17 of the International Covenant on Civil and Political Rights (ICCPR) and Article 8 of the European Convention on Human Rights (ECHR)). This relationship takes on added significance in the contemporary socio-technological context where 'a small number of companies have a window into most of our movements online'.[29] In seeking to safeguard the integrity of our private lives and personal relationships with others,[30] the right to privacy carves out intellectual 'breathing space' for the individual, which is essential for freedom of thought.[31] Again, since privacy tends to be a well-established right, the scope and contours of which have been outlined in a rich body of jurisprudence and scholarship,[32] in some cases, when violations of freedom of thought occur, legal commentators might automatically assume that the right to privacy is the right that is most relevant.

So understood, there are close associations, both conceptual and legal, between the right to freedom of thought on the one hand and the rights to freedom of expression and privacy on the other. However, a significant difference is that the latter rights are qualified, both in international human rights law and in most national legal orders. In contrast, the right to freedom of thought provides absolute protection to the internal process of thought prior to expression. Article 18(1) ICCPR and Article 9(1) ECHR, for instance, list freedom of 'thought' alongside 'conscience' and 'religion' but in Articles 18(3) ICCPR and Article 9(2) ECHR (the provisions that set out permissible limitations), only the 'freedom to *manifest* one's religion or beliefs' is subject to limitations, implying that freedom of (unmanifested) thought is an absolute right [emphasis added].[33] This interpretation of the these provisions is supported by the relevant *travaux préparatoires* and the argument that the right to freedom of thought is an absolute right is now the 'dominant view' in the literature.[34] For this reason, the right to freedom of thought has the potential to be a powerful tool when used in the appropriate context.

Before we conclude this section, it is important to highlight two more rights, the protective remit of which may overlap with that of the right to freedom of thought. The first is the absolute prohibition of torture or inhuman and degrading treatment or punishment (see Article 5 UDHR, Article 3 ECHR, and Article 7 ICCPR). As the Special Rapporteur notes, psychological torture can coercively alter or manipulate a victim's thoughts and thought processes.[35] Indeed, one concrete example in the ECHR's

travaux préparatoires of how the right to freedom of thought (alongside expression, conscience, religion and opinion) might be engaged is in relation to certain state practices of forced confessions and other 'abominable methods of police enquiry or judicial process which rob the suspected or accused person of control of his intellectual faculties and of his conscience'.[36] But, in practice, such examples of ill-treatment would fall under the absolute prohibition on torture, cruel, inhumane or degrading treatment or punishment, the remit of which extends to the psychological effects of ill-treatment,[37] or the right to a fair trial (see Article 10 UDHR, Article 14 ICCPR, and Article 6 ECHR). So understood, the prohibition on psychological torture might capture at least some cases that might otherwise be regarded as 'freedom of thought' cases.

Finally, while there is clearly conceptual overlap between the right to freedom of thought and the (qualified) right to mental integrity, as provided, for example, by Article 3(1) of the Charter of Fundamental Rights of the European Union, it is a relatively new right with comparatively little associated jurisprudence (at least compared to the more established rights to freedom of expression and privacy).[38] For this reason, it is difficult to argue that the existence of a right to mental integrity provides an explanation (even in part) as to why the right to freedom of thought has historically received so little attention.

3. Attributes of the right

3.1. Thought in the forum internum?

Although commentators sometimes use 'freedom of thought' interchangeably with 'freedom of expression' or as a label to categorise religious freedom and conscientious objection cases, freedom of thought is a freestanding right in international law, which can be separated from these related rights.[39] However, it is important to note that there is no legal authority and little scholarly analysis on what, precisely, is meant by the 'right to freedom of thought'. There is no consensus on what the right protects, or how it can be used. There is little clarity on what exactly 'thought' means in this context, or what 'freedom' means when applied to thought.

Despite these ambiguities, there are some clear starting points to help determine how the right might apply, if it has utility at all. Both Article 18 UDHR and Article 18(1) ICCPR list three rights, which are understood as distinct and 'equal': freedom of thought, freedom of conscience, and freedom of religion.[40] The relevant provisions of regional human rights treaties, such as Article 9(1) ECHR, adopt a similar approach. Each of these documents includes a separate but closely related right to freedom of expression. The decision to make explicit reference to 'thought' as distinct from 'conscience' and 'religion' was a deliberate one, as revealed in the *travaux préparatoires* of the UDHR. In his contributions, René Cassin, the French representative and one of the leading figures in introducing freedom of thought along with the Lebanese representative Charles Malik, emphasised the 'metaphysical significance' of 'freedom of inner thought', broadly understood.[41] Cassin argued that freedom of thought needed explicit protection because 'the opposite of inner freedom of thought was the *outward* obligation to profess a belief which was not held' meaning that the right could be worn away indirectly if it was not unconditionally protected.[42]

Against this background, 'thought' can be understood as encompassing all forms of *unmanifested* mental activity, including 'deliberation, imagination, belief, reflection, reasoning, cogitation, remembering, wishing, sensing, questioning, and desiring'.[43] However, not only may such a conclusion seem tautologous, important questions also arise about the desirability and consequences of a legal right to absolute protection for *all* unmanifested thoughts.[44] Moreover, the conceptual distinction between unmanifested and manifested thought can be critiqued (e.g. from an 'extended cognition' standpoint; see below).[45] Yet, it is clear from the text of Article 18(1) ICCPR and Article 9 ECHR that there is a legal distinction, unmanifested thought enjoying absolute legal protection. As soon as thought is manifested through speech or behaviour, it may become subject to lawful limitations as is the case with laws that provide for religious expression. However, a crucial problem with this distinction, as touched upon earlier, arises from thought that occurs outside of the head.

3.2. Thought in the forum externum?

Despite legal arguments that thought may be understood to be limited to unmanifested mental activity, philosophical and psychological perspectives raise questions over the scope of what should count as thought. It is striking that The Report begins with Rene Descartes' famous expression 'I think, therefore I am'. The ensuing 'Cartesian' character of The Report is something of a double-edged sword, especially with respect to the 'Conceptual Framework' section.[46] On the one hand, The Report is admirable in the way that it mirrors Descartes's own broad conception of the kinds of activity that count as 'thought'. Yet, both frame thought as that which takes place in the *forum internum*. This runs counter to several recent (and growing) trends in philosophy of mind and cognitive science which emphasise the external scaffolding on which thought constitutively depends. Thus, the scope of The Report may be somewhat limited, insofar as it will be challenging to delineate what would most appropriately fall under the scope of the right to freedom of thought, as a distinct right, as opposed to the right to privacy.

Descartes held a mind–body dualist position, according to which, unlike other physical objects and systems (including animals), humans were uniquely constituted by an additional immaterial thinking substance—a soul—whose essential characteristic or defining feature was thought (unlike the physical which is essentially extended in space).[47] But by the term 'thought,' Descartes intended to connote a very wide variety of mental states and processes—including beliefs, desires, intentions, memories, reasoning, dreaming, linguistic capacity, and emotion—all of which seem appropriately covered by the discussion in The Report. Similarly, Descartes wrestles with a distinction that The Report acknowledges in passing: the fact that the term 'thought' can refer to both a *process* and to the *product* of that process. This may seem obvious or trivial; 'freedom of thought' should be protected both insofar as it refers to the process of reasoning whereby we arrive at beliefs, desires etc., and also insofar is it refers to the beliefs, desires etc., that we end up holding as the products of those processes. But this is especially important, given the four attributes considered by The Report. Whilst freedom not to disclose one's thoughts, and freedom from punishment for one's thoughts, seem to connote the conception of thought as a product of a process (*qua* beliefs, opinions), the mention of an 'enabling environment' in The Report gives us

scope to broaden our considerations to include the act or process of thinking itself, which should also be similarly protected.

In terms of the 'enabling environment' for thought, one respect in which The Report may be somewhat limited—along Cartesian lines—is its focus on thoughts (and the act or process of thinking) as something of an inner sanctum or *forum internum* in contrast to a *forum externum*. This conception has been dubbed the 'Cartesian Sandwich', since it paints an input-processing-output picture whereby only the processing (i.e. the filling of the sandwich) counts as genuine thought, with the rest relegated to merely peripheral body and world.[48] But this account has encountered a great deal of critical scrutiny since the late 1990s, from advocates of the so-called 'extended' or 'embedded' mind hypothesis in philosophy of mind,[49] and the 'situated' approach to cognitive science.[50]

Briefly, these views—to varying degrees—take mental states and processes, or cognitive systems, to be essentially dependent on, or even constituted by aspects of the body or of the environment. Certain conditions must still be met for aspects of the environment to count as part of a cognitive process: Clark, for example, argues the content of the environmental scaffolding must be readily available and more or less automatically endorsed as it would be with ordinary thought.[51] But, on very general functionalist principles—which are widely endorsed elsewhere in philosophy of mind and cognitive science—the boundaries between mind and world are not so clear cut, possibly context dependent,[52] and certainly not neatly aligned with the physical boundaries of skin and skull. As Clark and Chalmers put it:

> 'If, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process. Cognitive processes ain't (all) in the head!'[53]

One need not use sci-fi examples of mind-uploading and brain-computer interfaces to illustrate this point; diaries, shopping and to-do lists, smart phones, one's fingers, and (in some cases) other people, may all—under the right circumstances—constitute both the products and the processes of thought, on Clark and Chalmers's account.

We need not debate the metaphysics of this claim here; it remains somewhat philosophically controversial (although increasingly less so, particularly if one reads it as a methodological claim about how thought processes should be studied). But in these circumstances, determining what exactly should be covered by a right to freedom of thought may be trickier than The Report supposes. Importantly, The Report does acknowledge some of these considerations. For example, in Paragraph 16 it is noted that extending the protection of freedom of thought to these kinds of extended cognition may not be necessary if the environmental scaffolding used already receives qualified protection under the right to privacy. But The Report could go further, because there are more recent philosophical discussions that tie questions about extended and situated cognition directly to some of the attributes of freedom of thought that The Report mentions. For example, concerning the freedom not to reveal or disclose one's thoughts, Carter and Palermos have argued that under certain conditions (i.e. precisely those which support cognitive extension) data breaches and computer compromises might constitute a kind of personal assault, such that (e.g.) forcibly searching someone's smart phone could constitute a violation of the freedom of thought in the relevant

sense.[54] Needless to say, contemporary technology, and the increasing way in which we 'offload' cognitive processes using it, provides more and more opportunities for this.

A similar point may be made concerning protection from impermissible alteration of one's thoughts. Carter argues that taking the extended mind hypothesis seriously gives rise to new possibilities for a kind of thought manipulation.[55] If a person's thoughts are dependent on, or constituted by, aspects of the environment, then the manipulation of those external scaffolds by a third party would count as an indirect way of manipulating the thoughts themselves. To the extent that this might constitute a novel kind of violation of the freedom of thought, this kind of case will also need to be considered.

In short, The Report is impressive in its breadth of considerations, and the fact that it even mentions the various forms of the extended mind hypothesis is both notable and welcome. The 'Cartesian' influence is also quite reasonable and helpful in some respects, but in other respects The Report could be 'extended' even further, especially given its otherwise forward-looking and technologically sympathetic character. Indeed, although the Report has the Cartesian rational actor as its main focus, much of the contemporary information economy is based on the insights of behavioural psychologists and data scientists who have been able to study and exploit the highly *ir*rational characteristics—often below the level of conscious awareness—that drive human behaviour. We return to this issue in section 5.2 below, where we consider the relationship between freedom of thought and the regulation of digital technologies.

4. Brief comments on the violation of the right in the context of mental health

The Report examines potential violations of freedom of thought in seven fields, including 'mental health' and 'existing and emerging technologies', such as 'neurotechnology'. The extent to which freedom of thought is threatened in these fields varies significantly and it is not possible to analyse each of these categories within the confines of this article. For this reason, we give some brief consideration of issues arising from one specific threat to freedom of thought highlighted by The Report, namely mental health.

In relation to 'mental health', The Report states that:

> 'Several stakeholders suggested that some tools for "treating" people with intellectual, cognitive or psychosocial disabilities are abused in ways that may violate freedom of thought. For example, psychotherapies, shock treatments, lobotomies and forced medication – some of which the medical community has denounced – reportedly have been used to coercively alter the thoughts of individuals, forcibly reveal thoughts (beyond legitimate therapeutic purposes), punish "inferred" thoughts, or even physically modify brains, in separate or cumulative violations of the freedom'.[56]

On the other hand, The Report also notes that 'for people with certain mental conditions, one individual submits that treatment for mental health is necessary for "restoring" one's freedom of thought (e.g. if one experiences delusions)'.[57]

Here we run into the problem of how we wish to characterise thought as 'free'. A conception common among some philosophers is that thought is free if it follows the dictates of reason. This emphasis on free thought being reasoned thought was stressed, for example, by Immanuel Kant. For Kant, free thought was 'the subjection of reason to no laws except those which it gives itself'. If there is 'lawlessness in thinking', says

Kant, 'the freedom to think will ultimately be forfeited'.[58] A prototypical example of unfree thought for Kant was enthusiasm. This was the phenomenon, common in the 17th and 18th centuries, of people claiming to receive knowledge directly from God.

Reaching a balance in this area is complex. For example, paranoid delusions can lead to suffering, harm, and even death, but the distinction between a 'delusion' and an unusual or eccentric belief is not always entirely clear.[59] Governance and oversight of involuntary treatment are vital if these issues are to be balanced in a way that protects rights, including freedom of thought.[60]

It is also worth noting that various agencies within the United Nations take radically different positions on this issue, with some opposing involuntary treatment in all circumstances, and others endorsing it and arguing that it protects rights in certain circumstances.[61] In particular, the UN Subcommittee on Prevention of Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment argues that withholding treatment from people who lack capacity to consent to it could constitute cruel, inhuman or degrading treatment or punishment.[62] Consistent with this, The Report notes that 'some civil society members campaign to minimize or abolish forced treatment for mental health conditions, while others emphasize that it remains necessary in limited circumstances'.[63] The precise nature of these 'limited circumstances', and the 'legitimate therapeutic purposes' to which the Special Rapporteur refers earlier,[64] requires constant redefinition as knowledge of the biological brain grows, human rights standards evolve, and legal precedents are developed.

There are, as the Special Rapporteur notes, examples of good practice, as 'several states worldwide have undertaken efforts to protect against coercion in the provision of mental health treatment'.[65] As the Special Rapporteur demonstrates, the construct of 'freedom of thought' is a useful prism through which to consider and evaluate such initiatives. Yet it remains a challenge to develop a reliable method to assess when there has been a violation of this right in the context of mental health. Simple self-report is one measure, yet there is the potential that the violation could impair the ability to report it. More consideration hence needs to be paid to the practicalities of assessing the presence of violations.

It is also important to carefully examine empirical claims made at the intersection of mental health, neuroscience and freedom of thought. The fMRI study referred to in section 2.1 above, and highlighted in The Report, involved inferring suicidal thoughts with 91% accuracy from fMRI data. Yet this data was obtained from just seventeen people with suicidal ideation and seventeen controls. The authors also noted that their study required 'highly cooperative and focused participants'.[66] It does not, therefore, provide a basis for clinical practice or firm conclusions. A similar caution applies to The Report's observation that, 'in 2019, several forensic psychiatrists claimed neuroimaging data could "feasibly" help to determine the likelihood of recidivism'.[67] Again, this study included a small number of participants, has various other limitations (as the study's authors note),[68] does not constitute a basis for firm conclusions, and raises ethical issues yet to be resolved.

For now, therefore, and for the foreseeable future, neuroimaging is still a considerable distance away from either predicting suicide or recidivism, diagnosing mental illness, or reading thoughts – a situation that is identical to the position more than a decade ago.[69] Wisely, the Special Rapporteur points out that 'while the capability of neurotechnology to reveal thought might be impressive within tightly controlled laboratory conditions, the

accuracy is far lower in the real-world, at present, and it is allegedly unable to passively "decode" thoughts that researchers have not predefined'.[70] This is a careful, balanced insight that is much needed in this field.

Overall, commentary on the field of neuroscience tends towards hyperbole, alarmism, and exaggeration of findings.[71] The Report is correct to highlight neurotechnology as an area of potential concern, but the relevant technology is still sufficiently primitive that there is adequate time to develop ethical standards to govern dilemmas which might arise when – and if – 'neurotechnology' develops to a point where it can decipher thoughts outside 'tightly controlled laboratory conditions.' It is far from that point at present.

A psychological or sociological view of this mismatch between what neuroscience is currently capable of and alarmist commentaries on the field of neuroscience, raises the question of whether societal responses to neurotechnology, big data, artificial intelligence, and machine learning, including our current considerations of their impact on freedom of thought, may have the characteristics of a moral panic. Moral panics involve a societal reaction to an alleged threat that is completely out of proportion to the actual harm being done or the risk of such harm. They are often based on the idea that 'people are motivated to act by mysterious and unrealistically powerful forces'.[72] In the 1950s, a moral panic emerged over the idea that reading gory comic books produced juvenile delinquency. In the 1960s, it was the effects of LSD. In the 1980s, a moral panic arose over satanic child abuse in day care centres. In the first decade of the 2000s, we saw a moral panic over the idea that violent video games lead to school shootings.[73] Goode and Ben–Yehuda identify three sources of moral panics: 1) grassroots public anxiety stemming from genuine public concerns; 2) elites seeking to benefit their own interests by creating a threat; and, 3) activist groups and moral entrepreneurs in the middle level of society.[74] Regulatory responses to neurotechnology, including questions about whether the right to freedom of thought has any practical utility in this arena, could be profitably analysed within this framework which would help guide targeted and effective law-making. This would be an example of how the use of a variety of disciplinary perspectives can shed light on current discussions of the right to freedom of thought.

5. Assessing the report's recommendations

5.1. Compliance with the right to freedom of thought

The Special Rapporteur acknowledges that despite Article 18 (1) ICCPR stating '[e]veryone shall have the right to freedom of thought, conscience and religion', the general recognition of the right to freedom of thought lags behind that of conscience and religion because it is 'underdeveloped in theory and practice'.[75] Noting that 'further clarity on the legal content and scope' of the right to freedom of thought is 'desirable', he encourages a General Comment to be adopted on the right, and for States, civil society and mental health professionals to take certain steps towards helping establish that clarity.

More specifically, the Special Rapporteur calls on States to 'review their legal and policy frameworks' to ensure that other rights, such as freedom of expression and privacy, which impact the right to freedom of thought, are also upheld in line with international human rights law.[76] In addition, States should hold public consultations with

members of vulnerable groups, human rights institutions, civil society, and technology companies to allow for full consideration of the necessary protections for people's *forum internum* freedoms, including the right to freedom of thought.[77] States should also engage the UN system to help clarify the right and support human rights defenders who are monitoring violations of the right.[78]

The Special Rapporteur recommends that members of civil society advocate for States to review and assess their legislation to increase compliance with international human rights law, including existing obligations that affect freedom of thought.[79] They could also provide critical thinking training, particularly for children, to discern and identify misinformation and/or disinformation.[80] Mental health professions should 'firmly establish human rights as core values when prioritising mental health interventions' including in relation to forced treatment.[81]

These recommendations are broad and perhaps can best be described as aspirational in nature. Given the lack of awareness of the right to freedom of thought as a freestanding right, it is difficult to gauge the likelihood of State actors taking the steps outlined above, not to mention whether there is any appetite for a new General Comment.[82] Moreover, in most countries, public consultations of the sort envisaged by the Special Rapporteur are the exception rather than the norm. If there are no public consultations on other more established rights, how likely is it that there will be public consultations on the relatively underdeveloped right that is freedom of thought?

While State actors may be slow to respond to the Special Rapporteur's call to action, non-State actors have a potentially pivotal role to play in developing the right to freedom of thought. Civil society organisations such as the Irish Council for Civil Liberties (ICCL), of which one of our authors is a member, could (i) map incidents where the right is not being upheld and highlight problematic policies and laws which interfere with the right, (ii) make specific calls on governments to safeguard interests in freedom of thought, (iii) take legal action to help establish the underdeveloped right, all of which results in (iv) greater awareness of the right.

Civil society organisations are already taking legal actions in respect of other related rights such as the right to privacy and the protection of personal data. Against this background, when it comes to the need for 'clarity on the legal content and scope' of the right to freedom of thought, at least some clarity may emerge in the context of such ongoing actions. This is because the argument can be made that the misuse of personal data has enabled technology firms to breach people's privacy rights and, by extension, undermine the right to freedom of thought.[83]

Consider some concrete examples. For many years, privacy advocates have argued against targeted advertising based on surveilling or tracking people's behaviour online.[84] Online advertising is primarily automatic, moves at speed, and happens behind the scenes, unlike traditional advertising. One of the main methods, real-time bidding (RTB), is an automated online auction system that auctions billions of advertisement slots every day. In simple language, when a person clicks on a website on their phone or laptop, for a split second there is a blank box where an advertisement will eventually appear. But in that split second, personal data about them and their online preferences are sent to multiple companies to solicit their bids for the opportunity to place their ad in the advertising slot.[85] ICCL research shows that, on average, a person in the US has their online activity and location exposed 747 times every day by the RTB industry,

while in Europe, RTB exposes people's data 376 times a day.[86] ICCL is taking action in European courts to end this misuse of personal data.[87] Not only does litigation of this sort provide a model for future legal actions in respect of the right to freedom of thought, viewing the issues at stake in ongoing privacy litigation through the prism of freedom of thought may help us better define the contours of the right and its relationship to privacy.

Another potentially important means to safeguard interests in freedom of thought is the regulation of technology companies. In this context, in the light of the Guiding Principles on Business and Human Rights: Implementing the United Nations 'Protect, Respect and Remedy' Framework, the Special Rapporteur makes several recommendations about the responsibilities of technology companies.[88] First, these companies should examine how their products or services might infringe the right to freedom of thought, paying particular attention to the effects on vulnerable people, including children.[89] Where problems are revealed, these companies should 'adopt alternatives' that are more human rights-compliant.[90] At the same time and on a regular basis, technology companies should publish transparency reports that set out the 'challenges' the companies face in complying with the right, as well as detailing their responses to these challenges.[91] In addition, digital platforms should 'facilitate independent research on compliance of their products and processes with international human rights law' e.g. by allowing independent bodies to carry out human rights impact assessments.[92] Finally, the Special Rapporteur makes specific mention of neurotechnology companies, recommending that they 'ensure a robust, privacy-focused and human rights-compliant framework for the collection, processing and storage of neurodata'.[93]

If adopted by technology companies, these strategies would undoubtedly make a significant contribution to the protection of freedom of thought. Protecting the rights of internet users, in relation to both privacy and freedom of thought is also likely to be beneficial to users' mental health.[94] However, given that the Guiding Principles on Business and Human Rights are soft law, how realistic is the prospect of technology companies adopting these strategies? Not only do the guidance principles lack an enforcement mechanism, but technology companies already raise concerns about the 'heavy burden' of complying with various regulations.[95]

5.2. Freedom of thought, the Digital Services Act and the AI Act

If the Special Rapporteur's recommendations are to be realised, what is required, we submit, is governance (by hard law) that provides for 'human rights-centered design, deliberation, and oversight' of technology.[96] Against this background, we now turn to consider two legislative initiatives on the part of the European Union, one a recently implemented Regulation (the Digital Services Act (DSA))[97] and the other a proposal, which is currently in the legislative process (the AI Act).[98] Both Acts, we argue, provide means for the right to freedom of thought to shape the governance of technology companies and implement many of the Special Rapporteur's recommendations.

The Digital Services Act, which came into force in November 2022, is a Regulation seeking to harmonise the approach of EU Member States in their efforts to create 'a safe, predictable and trusted online environment', one in which the fundamental rights set out in the Charter of Fundamental Rights of the European Union ('the Charter') are effectively safeguarded.[99] Among other things, the Act clarifies the rules on the

liability of providers of intermediary services in the face of user-generated illegal content and places new due diligence obligations on those providers in respect of how they tackle illegal content, risks to fundamental rights and broader societal risks (e.g. how disinformation might affect electoral processes). Of particular interest, from the perspective of the right to freedom of thought, is the Article 25(1) DSA prohibition of deceptive and manipulative practices commonly known as 'dark patterns', which are features on websites that are designed to induce users into making certain decisions that they may not have initially intended to make. Examples include hidden data protection controls or misleading consent requests. In this context, Recital 67 of the DSA gives some useful practical examples of the practices it considers impermissible. Yet, the requirement in Article 25(1) that providers shall not 'materially [distort] or [impair] the ability of the recipients of their service to make free and informed decisions' remains hard to operationalise. A right to freedom of thought that offers a principled account of what precisely constitutes impermissible influence on thought could offer some much-needed clarity and theoretical underpinning for legal requirements and prohibitions of this sort.

The Act also addresses the risks of manipulation associated with closely targeted personalised advertising. Article 26(3) prohibits providers from presenting advertisements to users based on profiling that draws on the special categories of personal data outlined in Article 9(1) of the General Data Protection Regulation (GDPR).[100] As Recital 69 of the DSA stresses, targeted advertisements based on such sensitive data can 'potentially appeal to [users'] vulnerabilities' and therefore 'have particularly serious negative effects.'[101]

It is in section 5 of the DSA, however, where we find the most promising potential means for the right to freedom of thought to play a role in shaping the governance of at least some forms of technology companies. This section sets out special obligations for 'very large online platforms and very large online search engines', meaning platforms and engines that have at least 45 million active users in the EU each month.[102] This most obviously includes the likes of Google, Twitter and Facebook. Among other obligations, Article 34 DSA requires that such providers must carry out regular risk assessments, seeking to identify and assess 'any systemic risks' arising from the design or operation of their services and systems, including their recommender and advertising systems. Significantly, 'systemic risks' include 'any actual or foreseeable negative effects for the exercise of fundamental rights'.[103] Though the provision goes on to list certain fundamental rights (such as privacy and freedom of expression) that have been traditionally understood as especially relevant in the online environment, the list is a non-exhaustive one meaning that, in appropriate cases, providers should seek to identify and assess any risks to the right to freedom of thought under Article 10 of the Charter and the right to mental integrity under Article 3. In addition, a further systemic risk of 'serious negative consequences to the person's physical and mental well-being' is expressly mentioned in Art 34(1)(b), which provides us with further grounds to argue that the right to mental integrity should, where appropriate, be central to risk assessments undertaken by these very large providers. Where systemic risks are identified, providers are required to address those risks by putting in place 'reasonable, proportionate and effective mitigation measures', examples of which are outlined in Article 35. The rest of the section outlines additional due diligence requirements for very large online platforms and search engines that align with some of the Special Rapporteur's recommendations. Chapter IV of the Act sets out rules relating to oversight, enforcement and penalties, the latter necessary where

providers of intermediary services infringe the Act. So understood, the DSA is likely to be considerably more effective than relying on any soft law alternatives.

Against this background, the DSA offers an opportunity, at least in the context of intermediary service providers operating in Europe, to realise many of the Special Rapporteur's recommendations about technology companies' responsibilities towards safeguarding the integrity of the *forum internum*. This offers at least one path towards better protection of the right to freedom of thought in the face of socio-technological change. But other paths are also beginning to emerge. Take, for instance, the EU Commission's proposal for a new AI Act, which, according to its Explanatory Memorandum, seeks to establish a 'human centric' approach to regulation to ensure that people 'trust' that AI technology is developed and used 'in a way that is safe and compliant with the law, including the respect of fundamental rights.'[104] The draft Regulation distinguishes between different types of AI practices based on risk to health and safety or fundamental rights. AI that creates 'unacceptable risk' will be prohibited whereas high risk AI systems will be subjected to onerous due diligence requirements.[105] While the Commission's Explanatory Memorandum and the draft Recitals make reference to impacts on 'fundamental rights' in general terms, some of the AI practices specifically mentioned in the draft would have especially significant impacts on the rights to freedom of thought and mental integrity.

For example, Article 5(1)(a) of the Commission's draft prohibits the use of AI systems that '[deploy] subliminal techniques beyond a person's consciousness in order to materially distort a person's behaviour in a manner that causes or is likely to cause that person or another person physical or psychological harm'. [106] Ever since a moral panic in 1957 over subliminal advertising,[107] both the public and many experts have come to view subliminal messaging as the prototypical form of impermissible influence. A well-developed right to freedom of thought could offer a clear explanation for why such practices are inconsistent with respect for fundamental rights. For example, if the right to freedom of thought stipulated that influences need to be accessible to the reasoned thought of the thinker, to promote autonomous decision making, subliminal influences would be deemed a violation of this right. Yet, subliminal techniques are low hanging fruit. We are here inveighing against a practice that nearly all would agree is inappropriate and which is rarely utilised in the contemporary environment.[108] What urgently needs to be addressed is what forms of conscious stimuli should to be deemed manipulative by the right to freedom of thought. We need to probe the grey areas.

Bearing on this point, the Explanatory Memorandum to the draft Act refers to '[o]ther manipulative or exploitative practices affecting adults' that may be facilitated by AI systems.[109] A well-developed right to freedom of thought could help clarify what should be taken to constitute 'manipulative or exploitative practices' (a point also applicable to the DSA). From a psychological perspective, this could involve the identification of key components of thought, such as attention, reasoning, and reflection, along with guidelines for how practices may impermissibly engage (or fail to engage) with these processes.[110] It is also worthy of note that Article 5(1)(b) of the draft AI Act specifically prohibits practices that '[exploit] any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, in order to materially distort the behaviour of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm'. Whilst this provision is clearly to be

welcomed, psychological research into human reasoning shows that we all possess cognitive and affective vulnerabilities, which AI has the potential to exploit.[111] As we mentioned above, at the end of section 3.2, the law needs to recognise that a 'reasonable person' and 'rational actor' may often be both unreasonable and irrational. As such, the insights of behavioural science must be incorporated into the design of the right to freedom of thought and related legislation.[112]

When the AI Act comes into force, the precise meaning and contours of its provisions will be worked out through litigation before the courts, among other fora. Given that the protection of fundamental rights outlined in the Charter is a core objective of the AI Act, there is, we submit, real and important scope for arguments about the rights to freedom of thought and mental integrity to feature in such litigation, inform how courts interpret the provisions and thereby shape the governance of AI in Europe.

6. Conclusions

The right to freedom of thought is something of a legal enigma. On the one hand, it features at the centre of international and domestic bills of rights, as an absolute right of seemingly profound importance to the individual and society. On the other hand, it has rarely featured in litigation to date and scholarly interest in the right remains limited.

Against this background, perhaps the main value of the Special Rapporteur's Report lies in the way it helps to renew our interest in this 'forgotten freedom',[113] orienting our attention to the possibilities that are latent within this right. For instance, one of the defining features of the right is the fact that it is not a qualified one, like most rights, but is absolute. Absolute rights are powerful instruments since they offer as complete a protection as can be offered in human rights law. But absolute rights can also be blunt instruments, and this perhaps explains why courts take such care in outlining the contours of such rights. In the case of the right to freedom of thought, a textual analysis of Article 18 of the ICCPR and corresponding provisions in regional human rights treaties, such as Article 9 of the ECHR, reveals a fundamental Cartesian distinction between unmanifested and manifested thought, the former absolutely protected and the latter subject to lawful limitations. But, as we have explored in this article, this seemingly clear legal distinction can be critiqued from philosophical and psychological perspectives. The extended mind thesis, for example, demonstrates that the Cartesian model is perhaps too simplistic and fails to capture the range of ways that thought is generated and developed.

The questions, then, as to what exactly constitutes 'thought' and what forms of thought enjoy absolute as opposed to qualified protection, cannot be answered by lawyers working in their own silo. This is also true of other difficult questions that emerge in the context of determining the contours of the right. For instance, what is the boundary between permissible influence on another person's thought and impermissible manipulation?[114] Dialogue between scholars in law and both philosophy and the mind sciences (psychology, psychiatry) is therefore required.[115] For such work to occur, extensive groundwork is likely to be needed to allow those with narrow disciplinary expertise in individual fields such as law, the mind sciences or philosophy, to gain a working understanding of the concepts and state of knowledge in these other fields. Psychiatrists, whose work often straddles law and the mind sciences, appear likely to be able to play a key role in helping bridge the knowledge gap between those with expertise in the

legal area and those with expertise in the mind sciences. Future empirical work in this area could include the use of the Delphi technique[116] (well-suited to areas where the state of knowledge is unclear)[117] to assess initial expert consensus in this area. There will also be the need for psychometrically valid measures of freedom of thought to support research in this area. At present, the only such tool of which we are aware, a 'Freedom of Thought' subscale of The Scales of Civil Rights questionnaire, [118] lacks face validity.

Achieving a more holistic understanding of the scope of the right to freedom of thought is essential if the right is to have more than just symbolic value. While many concerns about the threats of neuroscience technology are hyperbolic in nature, there are, for instance, real concerns about the impact of intrusive data collection and surveillance activities on our freedom of thought.[119] In this context, the Special Rapporteur's recommendations are to be warmly welcomed, even if some of these recommendations appear, at first glance, to be more aspirational than anything else. However, drawing on relevant sections of the EU's Digital Services Act, we have explored how some of the recommendations about the responsibilities of technology companies might be implemented. In addition, we have predicted that the rights to freedom of thought and mental integrity, as outlined in the Charter of the Fundamental Rights of the EU, have the potential to shape the governance of AI in Europe in the context of the legislative proposal for an AI Act. We argue that more human rights-informed approaches to the safeguarding of the right to freedom of thought are likely to emerge internationally once policy-makers reorient their attention towards that 'forgotten freedom'.

Notes

1. Ahmed Shaheed, *Interim report of the Special Rapporteur on freedom of religion or belief: Freedom of Thought*. A/76/380. United Nations, Geneva (2021). Retrieved from https://www.ohchr.org/sites/default/files/Documents/Issues/Religion/A_76_380_AUV.docx (accessed January 16, 2023)
2. Shaheed, Interim Report, [7].
3. See, for example, Sjors Ligthart and others, 'Rethinking the Right to Freedom of Thought: A Multidisciplinary Analysis', *Human Rights Law Review* 22, no. 4 (2022): ngac028.
4. Shaheed, Interim Report, [4].
5. Shaheed, Interim Report, [25].
6. Shaheed, Interim Report, [49]–[83].
7. Shaheed, Interim Report, Summary.
8. Karen Yeung, Andrew Howes and Ganna Pogrebna, 'AI Governance by Human Rights-Centered Design, Deliberation, and Oversight: An End to Ethics Washing', in *The Oxford Handbook of Ethics of AI*, ed. Markus D. Dubber, Frank Pasquale and Sunit Das (Oxford: OUP, 2020), 76–106.
9. UNECOSOC, Commission on Human Rights, Third Session, Summary Record of the Sixtieth Meeting (23 June 1948) UN Doc E/CN.4/SR.60, 10.
10. *Palko v. Connecticut*, 302 U.S. 319, 327, 58 S. Ct. 149, 82 L. Ed. 288 (1937)
11. Shaheed, Interim Report, [1].
12. Shaheed, Interim Report, [48].
13. For a more sceptical view of the right to freedom of thought as a 'distinctive virtue' or what it can offer over and above 'traditional understandings of personal autonomy or liberty', see Frederick Schauer, 'Freedom of Thought?', *Social Philosophy and Policy* 37, no. 2 (2020): 72–89.

14. Jan. C. Bublitz, 'Freedom of Thought in the Age of Neuroscience' *Archiv für Rechts- und Sozialphilosophie* 100, no. 1, (2014): 1–25.
15. G.L. Stelzer, 'Chemical Castration and the Right to Generate Ideas: Does the First Amendment Protect the Fantasies of Convicted Pedophiles', *Minnesota Law Review* 81, (1997): 1675–1709.
16. Susie Alegre, 'Rethinking freedom of thought for the 21st century', *European Human Rights Law Review* 3 (2017): 221–233; Susie Alegre, 'Opinion on the Implications for the Right to Freedom of Thought and the Right to Mental Integrity of data sharing by mental health websites in the UK and EU Member States as evidence in Privacy International's report – Your Mental Health for Sale' (2019): https://digitalfreedomfund.org/wp-content/uploads/2020/12/Susie-Alegre-Opinion-The-Right-to-Freedom-of-Thought-and-Data-Sharing-from-Mental-Health-Websites-Final.pdf (accessed January 16, 2023).
17. Ben. P. Vermeulen (2006), 'Freedom of Thought, Conscience and Religion (Article 9)', in *Theory and Practice of the European Convention on Human Rights*, ed. Pieter van Dijk and others, 4th ed (Antwerp: Intersentia, 2006), 752.
18. Jan. C. Bublitz, 'The Nascent Right to Psychological Integrity and Mental Self- Determination', in *The Cambridge Handbook of New Human Rights*, ed. Andreas von Arnauld, Kerstin von der Decken and Mart Susi (Cambridge: Cambridge University Press, 2020) 387–403, 395.
19. Shaheed, Interim Report, [76–78].
20. Shaheed, Interim Report, [77–78].
21. See Adrienne Stone and Frederick Schauer, eds, *The Oxford Handbook of Freedom of Speech* (Oxford: Oxford University Press, 2021).
22. Shaheed, Interim Report, [18].
23. For example, Article 10(1) ECHR explicitly states that the right to freedom of expression includes the freedom 'to receive and impart information and ideas'. See also *Autronic AG v Switzerland* [1990] ECHR 12726/87; *Mustafa and Tarzibachi v Sweden* [2008] ECHR 23883/06.
24. Jan. C. Bublitz, 'Freedom of Thought as an International Human Right: Elements of a Theory of a Living Right', in *The Law and Ethics of Freedom of Thought*, ed. Marc J. Blitz and Jan C. Bublitz (New York, NY: Palgrave Macmillan, 2021) 49–101.
25. Bublitz, 'Freedom of Thought in the Age of Neuroscience', p. 21.
26. Shaheed, Interim Report, [91].
27. L.S. Vygotsky, *Mind in Society: The Development of Higher Mental Processes*, Ed. Michael Cole and others (Cambridge, MA: Harvard University Press, 1978).
28. Charles Fernyhough, 'The Dialogic Mind: A Dialogic Approach to the Higher Mental Functions' *New Ideas in Psychology* 14, no. 1 (1996): 47–62.
29. Chris J. Hoofnagle and others, 'Behavioral Advertising: The Offer You Cannot Refuse' *Harvard Law & Policy Review*, 6 (2012): 273–296, 279.
30. In interpreting the remit of Article 8 ECHR, the ECtHR has identified a 'right to personal development' *and* a 'right to establish and develop relationships with other human beings in the outside world.' See *Pretty* v *United Kingdom* [2002] ECHR 2346/02 at [61]; *Oleksander Volkov* v *Ukraine* [2013] ECHR 21722/11 at [165].
31. See Neil M. Richards, *Intellectual Privacy* (Oxford: Oxford University Press, 2015).
32. See András Koltay and Paul Wragg, ed., *Comparative Privacy and Defamation* (Cheltenham: Edward Elgar, 2020); Gloria González Fuster, Rosamunde Van Brakel and Paul De Hert, ed., *Research Handbook on Privacy and Data Protection Law* (Cheltenham: Edward Elgar, 2022).
33. *Kokkinakis v Greece* (1993) Series A no 260-A [33]; Patrick O'Callaghan and Bethany Shiner, 'The Right to Freedom of Thought in the European Convention on Human Rights' *European Journal of Comparative Law and Governance* 8, no. 2–3 (2021): 112–145.
34. Ligthart and others, 'Rethinking the Right to Freedom of Thought', 10.
35. Shaheed, Interim Report, [49–50].

36. Pierre-Henri, Teitgen, 'Establishment of a Collective Guarantee of Essential Freedoms and Fundamental Rights', Report to the Committee on Legal Affairs and Human Rights, 5 September 1949, Doc 77
37. Shaheed, Interim Report, [49–51].
38. Some scholars are beginning to explore the overlap between the rights to freedom of thought and mental integrity. See, for example, Ligthart and others, 'Rethinking the Right to Freedom of Thought', 6–8
39. See O'Callaghan and Shiner, 'The Right to Freedom of Thought in the European Convention on Human Rights'.
40. General comment No. 22 (48) (art. 18) of the Human Rights Committee (27 September 1993) at [1]. See also Shaheed, Interim Report, [1] and [22].
41. UNECOSOC, Commission on Human Rights, Third Session, Summary Record of the Sixtieth Meeting (23 June 1948) UN Doc E/CN.4/SR.60, 10 and 13.
42. UNECOSOC, Commission on Human Rights, Third Session, Summary Record of the Sixtieth Meeting, 10.
43. Lucas Swaine, 'Freedom of Thought as a Basic Liberty' *Political Theory* 46, no. 3 (2018): 405–425, 411.
44. Ligthart and others, 'Rethinking the Right to Freedom of Thought', 3.
45. Shaheed, Interim Report, [16].
46. Shaheed, Interim Report, [10–16].
47. See, for example, Rene Descartes *Meditations on First Philosophy* In *The Philosophical Writings of Descartes, Vol. II*, edited by J. Cottingham, R. Stoothoff, and D. Murdoch (Cambridge: Cambridge University Press, 1984)
48. Susan Hurley, 'Perception and Action: Alternative Views' *Synthese* 129 (2001): 3–40.
49. Andy Clark and David Chalmers, 'The Extended Mind' *Analysis* 58, no. 1, (1998): 7–19; Anthony Chemero, *Radical Embodied Cognitive Science*, (Cambridge, MA: MIT Press, 2010).
50. Eliot R. Smith and Elizabeth C. Collins, 'Situated Cognition' in *The Mind in Context*, ed., Batja Mesquita, Lisa Feldman Barrett and Eliot. R. Smith (New York, NY: Guildford Press, 2010) 126–145; Joel Walmsley 'Methodological Situatedness; or, DEEDS worth Doing and Pursuing' *Cognitive Systems Research* 9, no. 1–2 (2007): 150–159.
51. Andy Clark, 'Coupling, Constitution and the Cognitive Kind: A Reply to Adams and Aizawa', *The Extended Mind*, ed., Richard Menary (Cambridge, MA: MIT Press, 2010) 81–99.
52. John Haugeland, 'Mind Embodied and Embedded, in *Having Thought: Essays in the Metaphysics of Mind* (Cambridge, MA: Harvard University Press, 1998) 207–237.
53. Clark and Chalmers, 'The Extended Mind', p. 8.
54. J. Adam Carter and S. Orestis Palermos, 'Is Having Your Computer Compromised Personal Assault? The Ethics of Extended Cognition', *Journal of the American Philosophical Association*, 2 (2016): 542–560.
55. J. Adam Carter, 'Varieties of (Extended) Thought Manipulation' in *The Law and Ethics of Freedom of Thought*, ed. Marc J. Blitz and Jan C. Bublitz (New York, NY: Palgrave Macmillan, 2021) 291–309.
56. Shaheed, Interim Report, [80].
57. Shaheed, Interim Report, [81].
58. Immanuel Kant, 'What Does It Mean to Orient Oneself in Thinking?' in *Religion within the Boundaries of Mere Reason and Other Writings*, eds., Allen Wood and George di Giovanni (Cambridge: Cambridge University Press, 1998), p 13.
59. Cf *Mockutė v Lithuania* [2018] ECHR 66490/09.
60. Cf Fleur Beaupert, 'Freedom of Opinion and Expression: From the Perspective of Psychosocial Disability and Madness', *Laws* 7, no. 1 (2018): 3.
61. B. D. Kelly, 'Mental capacity, human rights, and the UN's Convention on the Rights of Persons with Disabilities', *Journal of the American Academy of Psychiatry and the Law*, 49 (2021): 152–156.

62. United Nations Subcommittee on Prevention of Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, 'Approach of the Subcommittee on Prevention of Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment regarding the rights of persons institutionalized and treated medically without informed consent' (2016) UN Doc CAT/OP/27/2. New York: United Nations.
63. Shaheed, Interim Report, [82].
64. Shaheed, Interim Report, [80].
65. Shaheed, Interim Report, [85]; S.P. Sashidharan, Roberto Mezzina and Danius Puras, 'Reducing Coercion in Mental Healthcare', *Epidemiology and Psychiatric Sciences* 28 (2019): 605–612.
66. Marcel A. Just and others, 'Machine learning of neural representations of suicide and emotion concepts identifies suicidal youth' *Nature Human Behaviour* 1, no. 12, (2017): 911–919, p. 917.
67. Shaheed, Interim Report, [77].
68. C. Delfin and others, 'Prediction of recidivism in a long-term follow-up of forensic psychiatric patients: Incremental effects of neuroimaging data'. *PLoS One*, 14, no. 5 (2019), e0217127.
69. Brendan D. Kelly, 'Brain imaging in clinical psychiatry: Why?' in *I know what you're thinking: Brain Imaging and Mental Privacy*, ed. Sarah D. Richmond, Geraint Rees and Sarah J. L. Edwards (Oxford: Oxford University Press, 2012) 111–122; Brendan D. Kelly, *In Search of Madness: a Psychiatrist's Travels through the History of Mental Illness* (Dublin: Gill Books, 2022).
70. Shaheed, Interim Report, [76].
71. Sally Satel, and Scott O. Lilienfeld, *Brainwashed: The Seductive Appeal of Mindless Neuroscience* (New York: Basic Civitas Books: 2013); Kelly, *In Search of Madness*.
72. Erich Goode and Nachman Ben-Yehuda, 'Grounding and Defending the Sociology of Moral Panic', in *Moral Panic and the Politics of Anxiety*, ed. Sean Hier (London: Routledge, 2011) 20–36.
73. E. Goode and N. Ben-Yehuda, *Moral Panics: The Social Construction of Deviance* (New York: John Wiley & Sons, 2010).
74. E. Goode and N. Ben-Yehuda, *Moral Panics*, p. 126.
75. Shaheed, Interim Report, [96].
76. Shaheed, Interim Report, [97].
77. Shaheed, Interim Report, [97].
78. Shaheed, Interim Report, [97].
79. Shaheed, Interim Report, [98].
80. Shaheed, Interim Report, [98].
81. Shaheed, Interim Report, [99].
82. We should note here, however, that the Advisory Committee to the UN Human Rights Council has recently received a mandate to prepare a report on 'neurorights'. Such a report could be a precursor to a General Comment in this area. See Resolution A/HRC/51/L.3 on Neurotechnology and Human Rights (29 September 2022).
83. Johnny Ryan, '*The Biggest Data Breach: ICCL report on the scale of Real-Time Bidding data broadcasts in the U.S. and Europe*' Irish Council for Civil Liberties, May 16, 2022. https://www.iccl.ie/digital-data/iccl-report-on-the-scale-of-real-time-bidding-data-broadcasts-in-the-u-s-and-europe/ (accessed January 16, 2023). See also O'Callaghan and Shiner, 'The Right to Freedom of Thought in the European Convention on Human Rights'.
84. David C. Gray and Stephen Henderson, ed., *The Cambridge Handbook of Surveillance Law* (New York: Cambridge University Press, 2017).
85. Ryan, 'The Biggest Data Breach'.
86. Ryan, 'The Biggest Data Breach'.
87. At Landgericht Hamburg against the tracking industry standards body IAB TechLab, and against Microsoft's online advertising exchange Xandr https://www.iccl.ie/rtb-june-2021/; at the Irish High Court against the Data Protection Commission, for its failure to investigate

our complaint about Google's RTB data breach https://www.iccl.ie/news/iccl-sues-dpc-over-failure-to-act-on-massive-google-data-breach/; and at the Brussels Market Court, ICCL is a party against IAB Europe's appeal of an order by the Belgian Data Protection Authority, and 27 other EU data protection authorities, against IAB Europe's 'TCF' consent spam system. https://www.iccl.ie/news/gdpr-enforcer-rules-that-iab-europes-consent-popups-are-unlawful/ (accessed January 16, 2023).

88. Shaheed, Interim Report, [100].
89. Shaheed, Interim Report, [100].
90. Shaheed, Interim Report, [100].
91. Shaheed, Interim Report, [100].
92. Shaheed, Interim Report, [100].
93. Shaheed, Interim Report, [100].
94. Elias Aboujaoude, 'Protecting Privacy to Protect Mental Health: the New Ethical Imperative'. *Journal of Medical Ethics* 45, no. 9 (2019): 604–607; Victoria Betton, V and James Woollard, *Teen Mental Health in an Online World: Supporting Young People around their Use of Social Media, Apps, Gaming, Texting and the Rest* (London: Jessica Kingsley Publishers. 2018).
95. In a 2021 consultation on the Digital Services Act and Digital Markets Act, organised by the Irish Department of Enterprise, Trade and Employment, Facebook made reference to risks the 'heavy burden' on 'Very Large Platforms' under the DSA. https://enterprise.gov.ie/en/Consultations/Consultations-files/Facebook-DSA-Submission.pdf (accessed January 16, 2023).
96. Yeung, Howes and Pogrebna, 'AI Governance by Human Rights-Centered Design, Deliberation, and Oversight'.
97. Digital Services Act, Regulation (EU) 2022/2065 (19.10.2022).
98. AI Act, Proposal of the European Commission, COM (2021) 206 final (21.04.2021).
99. Article 1(1) DSA.
100. These include 'personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, [and] genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation.' See Article 9(1) of the General Data Protection Regulation, Regulation (EU) 2016/679 (27.04.2016).
101. For minors, there is a heightened concern about the risk of such advertisements in exploiting vulnerabilities. Article 28(2) DSA prohibits advertisements based on profiling using 'personal data' (as opposed to 'special categories' of personal data) when providers 'are aware with reasonable certainty that the recipient of the service is a minor.'
102. Article 33 DSA.
103. Art 34(1)(b) DSA.
104. Section 1.1 of Explanatory Memorandum.
105. Article 5 outlines prohibited AI systems while Title III, Chapter 2 sets out requirements for high-risk AI systems.
106. Note that as we were completing this paper, the Council of the EU proposed the following wording for Article 5(1)(a) in its 'common position' or 'general approach' to the Act: 'the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness with the objective to or the effect of materially distorting a person's behaviour in a manner that causes or is reasonably likely to cause that person or another person physical or psychological harm'. (Brussels, 25.11.2022; 2021/0106 (COD)).
107. See William M. O'Barr, '"Subliminal" Advertising', *Advertising & Society Review* 6, no. 4 (2005); Ronald A. Fullerton, R. A. (2010). '"A Virtual Social H-bomb": the Late 1950s Controversy over Subliminal Advertising', *Journal of Historical Research in Marketing* 2 (2010): 166–173.
108. Indeed, in many European countries broadcasting codes and advertising regulations already prohibit subliminal advertising.

109. Section 5.2.2 of the Explanatory Memorandum.
110. Simon McCarthy-Jones, *The Battle for Thought* (London: Oneworld Publications, forthcoming).
111. Daniel Kahneman, Thinking, Fast and Slow (New York, NY: Macmillan, 2011).
112. We thank an earlier anonymous reviewer of this paper for stressing this point to us.
113. Dwight Newman, 'Interpreting Freedom of Thought in the Canadian Charter of Rights and Freedoms', Supreme Court Law Review 91 (2019): 107-122.
114. Jan C. Bublitz and Reinhard Merkel, Crimes against Minds: On Mental Manipulations, Harms and a Human Right to Mental Self-Determination', *Criminal Law and Philosophy* 8, no. 1 (2014): 51–77.
115. Simon McCarthy-Jones, 'Freedom of Thought: Who, What, and Why?', in *The Law and Ethics of Freedom of Thought*, ed. Marc J. Blitz and Jan C. Bublitz (New York, NY: Palgrave Macmillan, 2021) 27–47.
116. This involves a series of rounds of anonymous experts responding to sets of statements, with each round focussing on an iteratively refined set of materials. This process allows participants to reflect on their position in light of the views of others, yet avoids direct confrontation, making the process "more conducive to independent thought on the part of the experts to aid them in the gradual formation of a considered opinion". Norman Dalkey and Olaf Helmer, 'An Experimental Application of the Delphi Method to the Use of Experts', *Management Science* 9 (1963): 458–67. doi:10.1287/mnsc.9.3.458, p 459.
117. Marlen Niederberger and Julia Spranger, 'Delphi Technique in Health Sciences: A Map' *Front. Public Health* 8:457 (2020): doi: 10.3389/fpubh.2020.00457
118. Üzeyir Ok, 'Roles of Fundamentalism and Authoritarianism in Relations between Religiosity and Civil Liberties among Muslims', *Asian Journal of Social Psychology* (forthcoming): https://doi.org/10.1111/ajsp.12553
119. Alegre, 'Rethinking freedom of thought for the 21st century'.

Acknowledgements

The authors are grateful to Liam Herrick (Executive Director, Irish Council for Civil Liberties) and Professor Maria Cahill (University College Cork), for their support of this project and discussions. The authors are also grateful to Dr Conor Linehan for his participation in the interdisciplinary symposium in January 2022, from which this paper stemmed.

Disclosure statement

In June 2021, Patrick O'Callaghan, Bethany Shiner and Simon McCarthy-Jones participated in a workshop meeting with the Special Rapporteur on freedom of religion or belief, which supported the work of the Special Rapporteur in developing his report. No other potential conflicts of interest were reported by the authors.

Funding

This work was supported by a New Foundations 2020 grant from the Irish Research Council. In addition, Patrick O'Callaghan is funded by an Irish Research Council (Consolidator) Laureate grant (IRCLA/2022/2628).

Notes on contributors

Patrick O'Callaghan is a lecturer in law at University College Cork, Ireland. He is Principal Investigator of the Law and the Inner Self project (2022-2026), funded by an Irish Research Council (Consolidator) Laureate Grant.

Olga Cronin is a Senior Policy Officer in the area of Surveillance and Human Rights at the Irish Council for Civil Liberties (ICCL).

Brendan D. Kelly is Professor of Psychiatry at Trinity College Dublin and Consultant Psychiatrist at Tallaght University Hospital, Dublin. His research interests include mental health legislation, human rights, and the history of psychiatry.

Bethany Shiner is Senior Lecturer in Law at Middlesex University London and is completing her PhD at the University of Oxford on the scope and application of the right to freedom of thought under Article 9 of the European Convention on Human Rights.

Joel Walmsley is a lecturer in Department of Philosophy at University College Cork. His research interests are in philosophy of cognitive science, philosophy of mind, and Artificial Intelligence.

Simon McCarthy-Jones is Associate Professor in Clinical Psychology and Neuropsychology at Trinity College Dublin. His research interests include hallucinations, child sexual abuse, spite, and the right to freedom of thought.