UNIVERSITY OF CALIFORNIA

Los Angeles

Mental Privacy

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Philosophy

by

Jordan James Wallace-Wolf

2020

ABSTRACT OF THE DISSERTATION

Mental Privacy

by

Jordan James Wallace-Wolf

Doctor of Philosophy in Philosophy

University of California, Los Angeles, 2020

Professor Seana Shiffrin, Chair

When we catch someone reading our diary, we know what to say: "mind your own business!" But why is this the right response? What makes our diary our business, such that the other person should not mind it? And why does reading what we wrote down count as minding our business anyway? Of course, there might be embarrassing information in what we wrote, and some entries might leave us vulnerable to exploitation or manipulation, but the contents might also be innocuous, such as what we ate that day or what happened to us. Nonetheless, the colloquial rebuke feels right.

It feels right, we might hypothesize, because the reading itself is a violation of our privacy, and one that is accomplished, apparently, just by revealing information about what we think. This dissertation investigates and defends this hypothesis. It seeks to show how our thoughts are our business, how others revealing them interferes with it, and how preventing this kind of interference is an important goal of various moral, legal, and democratic norms.

Chapters I and II begin answering these questions as a pair. Chapter I argues that persons have a morally significant interest in thinking generally, which is achieved through a proper combination of thinking by oneself and thinking with others. The proper combination is partly determined by readiness or maturity. Thinking with someone when they are not ready is disruptive, and since thinking with someone is begun by finding out what she thinks, there are limits on the latter.

One such limit prohibits learning the thoughts that a person holds back in diary, but another prohibits learning the identity of a person who puts her thoughts forward anonymously. This kind of limit is the subject of Chapter II, as it plays out in anonymous reading and speech, and in voting.

Chapter III and IV are a paired examination of the status of thoughts in criminal law. Chapter III defends the thought crime doctrine, according to which persons may not be punished for their thoughts, and Chapter IV gives a partial defense of the privilege against self-incrimination.

The dissertation of Jordan James Wallace-Wolf is approved.

Barbara Herman

Alexander Jacob Julius

Richard M. Re

Seana Shiffrin, Committee Chair

University of California, Los Angeles

2020

To all the believers

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Around the end of 2011, I spoke with Seana Shiffrin on the phone to try to determine if I should study law and philosophy at UCLA. Even now, I remember coming away from that conversation with two thoughts: Seana was serious about these subjects and that if I wanted to study them, I should study them with her. I eventually decided that I was ready for a long academic journey, and I came to UCLA. Luckily for me, my impression of Seana was dead on. I found again and again that she was a passionate and consummate theorist of law and philosophy and that I was beyond fortunate to have her as an advisor. What I didn't know in 2011 was how much more I would learn from Seana, over and above scholarly tradecraft. As her advisee, I have learned how to better occupy the role of philosopher, teacher, colleague, and member of the moral community, because she is all these things, and exemplarily so. She pushed me, at some personal cost, to make this dissertation as good as it can be, and for that I am deeply grateful.

I got to know Barbara Herman in the seminar that she taught in my first year at UCLA. I enjoyed it immensely. So much so that, in retrospect, I think I did a poor job of controlling my enthusiasm, as evidenced by Barbara's comment to me one day at lunch, to the effect of "you like to talk don't you?" Since then, Barbara has patiently advised me on a number of projects, and I have always benefitted from her example. Characterizing her impact on me fully would be impossible, but I think what has most struck me is her combination of intellectual self-possession backed by an inexhaustible desire to learn, joined with a keen sensitivity to moral experience as it is lived by those like her, who are most invested in it.

My first memory of AJ was during a seminar he taught. I felt that the material was exciting and challenging. More seminars followed, and my original feeling never went away, and indeed, grew in my subsequent conversations with him. He has a remarkable ability to immediately meet someone

else's thinking, wherever it is, and to elevate it. At some points in this dissertation, I thought that I better understood some of his ideas that had most puzzled me before. His influence is throughout.

I have Richard Re to thank for the subject matter of this dissertation. In the spring of 2016, I took his seminar about the Fourth Amendment. The questions it tackled were profound, and not least because of his way of posing them. Richard approached the subject matter with a contagious spirit of excitement and curiosity. His methodology was genuinely diverse, making use of history, technology, law, philosophy, and more. Afterward, he advised me on a standalone paper on some of the themes of the course. While writing that paper, I decided that I wanted to continue with a dissertation-scale investigation and that I would need his insight to do it. He generously came on board.

Outside of my committee, several other faculty members played a critical role in developing my ideas and offering comments on drafts. Pamela Hieronymi helped me throughout my time at UCLA, including fielding phone calls about her work, reading a number of drafts of mine, and offering her advice about how to be a philosopher. Her work about the wrong kinds of reasons has been especially influential, and the clarity of her writing and thinking is an ideal. Mark Greenberg helped me at a number of points and taught me to be more careful. Larry Sager also provided a fresh pair of eyes at a critical moment. Thanks also to Gavin Lawrence and Andrew Hsu who both kindly opened their doors to me.

I would also like to thank a number of former and current UCLA graduate students, whose passion and dedication inspired me, helped me to think better, and some days, to feel better too. Thanks to Jonathan Gingerich, Sabine Tsuruda, and Brian Hutler, for helping me find my footing at UCLA and in the study of law. Thanks to Bill Kowalsky, Gabe Dupre, and Gabby Johnson, for rolling up their sleeves and engaging with my ideas anytime, anywhere. Thanks to Jenna Donohue,

Amber Kavka-Warren, and Andrew Flynn for commenting on and proofreading this dissertation. The errors are all mine, but there are fewer because of them. Thanks especially to Jenna for her support, both as a philosopher and as a parent. Thanks especially to Amber for her aggressive Bluebooking and thoughtful comments across multiple drafts. Thanks especially to Andrew for his kindness, humor, and deep thinking. Thanks to Brad McHose for his collegial spirit and fierce questions. Thanks to Jeff Helmreich for meeting with me at an early stage in my thinking. Thanks to Michael Noltemeyer for last minute edits. Thanks to the participants of EWS and to the 2019-2020 legal writing seminar for feedback on the ideas in this dissertation.

Many people helped keep the lights on while I wrote this dissertation. Thanks to my parents for encouraging me, and for hosting me and my family during a pandemic. Thanks to Wally and Ceciliee Wallace for doing the same. Special thanks to my mom and Ceciliee for doing so much to care for my family while I wrote. It's your dissertation too. Thanks to the daycare workers at Bright Horizon Westwood, for caring for Rosie in hard times, and thanks to Doug Myers, who went above and beyond to make sure I checked all the boxes.

Thanks most of all to my girls, who made it all worthwhile: Stephanie, Rosie, and JJ. Through you, I got insight when I was stuck, love when I was down, and another chance when I made a mistake. You believed in me.

VITA

EDUCATION

Yale University
B.A., *Magna Cum Laude* 2008

Tufts University
M.A. 2010

University of California Law School
J.D., *Order of the Coif* 2017

EMPLOYMENT

University of California, Los Angeles
Teaching Associate 2012-2014
Teaching Assistant 2014-2019

PRESENTATIONS

"The Thought Crime Doctrine
and Limits of Criminal Law"
Ethics Writing Workshop, Philosophy Department
University of California, Los Angeles Fall 2019

"The Privacy of Thought"
Ethics Writing Workshop, Philosophy Department
University of California, Los Angeles Fall 2020

Introduction

Colloquially, a certain kind of privacy violation is rebuked with "mind your own business!" which, shorn of its commercial, anglophone bluntness, translates as "pay attention to your own affairs!" Both phrases suggest that persons have things that are their concern, their responsibility, or for them to handle, such that others act wrongly by getting mixed up in them. The wrong here is one of usurpation or officiousness – of a sphere or domain being crossed into by another.

Examples are prevalent, but their conceptual scope and moral core are obscure. For instance, trying to mediate an argument between two strangers on a park bench, or disciplining a stranger's child will both earn a swift "mind your own business!" The arguing or the disciplining is, within limits, for them to conduct. A different kind of example concerns facts rather than activities. It's a violation of informational privacy to look into someone's sexual orientation, or to access a person's medical, financial, educational, or criminal record(s). Snooping in someone's medicine cabinet, for example, will earn an emphatic "mind your own business!" Why are these activities and these facts part of a person's domain of concern? And where is the limit of that domain? After all, others may need to intervene in disputes and in child-rearing, and they will often need information about things like one's finances or criminal background.

The latter kind of example, about accessing information, is puzzling for an additional reason: it is unclear whether, and if so how, wrongfully accessing information is a distinct kind of wrong. No doubt, sensitive information about a person may make her vulnerable to exploitation or discrimination, but if concerns about exploitation and discrimination are the dominant concerns about privacy violations, then it seems that violating someone's privacy is not a distinct wrong, but wrong only if and because it is a part of exploiting or discriminating. Similarly, a person may feel

shame or embarrassment if they are aware of an information privacy violation, but again, it seems that the latter is wrong only if and because it is a part of gratuitous shaming. And yet it seems that it is a violation of informational privacy to learn facts about a person when they will not be used to discriminate or exploit, and when the target does not know about it, and so is not shamed or embarrassed.

This dissertation investigates these questions about a person's domain of concern, not by trying to survey it all at once, but by narrowing in on one aspect of it, that of mental activity. This choice of focus is warranted by the wide scale social changes being wrought by information technology and invited by the feeling, found in law, history, and literature, that there is something distinctly troubling about peering into a person's mind.

The result of this investigation is a moral claim: that a critical part of a person's business or domain of concern is thinking, and that this activity has a distinct value that makes interfering with it, including revealing information about it, wrongful in a correspondingly distinct way. The argument for this takes place across four chapters organized into two pairs (Chapters 1 & 2 and Chapters 3 & 4), with the paired chapters being companion discussions of a larger theme.

Chapters 1 and 2 are about the harms of exposing information about a person's thinking. Chapter 1 lays the framework for this topic. Its main claim is that thinking is a core part of a person's business and that conducting it well involves being able to postpone thinking with others until one has sufficiently readied oneself for it, by thinking first without them. This claim is paired with a second, which is that thinking with another is partly constituted by another mind being aware of what one thinks. Together, these claims lead to the conclusion that finding out what someone thinks without waiting for her to make it known risks precipitously converting her solitary thinking

into joint thinking. This risk may not be run, because catching a thinker unprepared endangers her thinking, in the way that showing up early to someone's party disrupts her hosting.

Chapter 2 approaches this same argument from a different direction. It takes on the assumption that thinkers should be able to defer joint thinking until they are ready, but it argues that this deferral may be accomplished not just by withholding one's thoughts, but also by making them known under cover of anonymity, i.e., not as one's own. In this way, persons may justifiably avoid the harms of having their thoughts exposed, while still participating in valuable communal thinking, such as by reading public information and speaking to the public, anonymously. I apply this participatory view of anonymity to three democratic practices: the secret ballot, ballot initiatives, and voting by legislative representatives.

Chapters 3 and 4 argue that the value of thinking, and the need for persons to think by themselves before thinking with others, set limits on the rightful scope of criminal law. Chapter 3 focuses on the *ex ante* perspective, i.e., on how far criminal law can go in preventing crime. According to the thought crime doctrine, criminal law can extend no further than the law of attempts. Neither a belief that a crime should be committed nor the intention to commit it may be punished, though the first step toward actually committing it may be. I defend this doctrine as necessary to protect the opportunity for persons to think about what to do, even though it has the serious cost of precluding the apprehension and punishment of those who are set on committing crimes.

Chapter 4 takes up the same themes, but from the *ex post* perspective, i.e., how far into someone's mind the criminal law may go in order to convict persons of the crimes they commit. I argue that just as persons have an interest in thinking about what to do, they have an interest in thinking about what they have done. This latter kind of thinking is threatened by compelled

confession, whether accomplished through the threat of contempt or by a mind-reading machine, and so provides some principled support for the Fifth Amendment privilege against self-incrimination, though whether the privilege is justified all things considered depends on other moral and legal values.

The view that emerges is one of mental privacy. A person's business includes thinking, and carrying it out requires that others not "mind" it in various ways, such as by prematurely revealing its contents and thereby transforming its character from solitary to joint, or by burdening it with punishment or other consequences that distort its direction or halt it altogether.

Chapter 1: Accessing Thoughts

I. Introduction

Information and communication technologies are deeply changing forms of social life. That much is a cliché, but it is one that bears repeating, because the moral implications of these changes are not well understood. Our collective power to gather and organize information is rapidly outpacing our collective understanding of how that power should be wielded.

One prominent family of norms governing the use of information is information privacy. Many feel that it places moral limits on how some information can be gathered and shared, but it's hard to take seriously without a good grasp of why we need it,[1] and it's hard to get a good grasp of why we need it because it appears to do, by design, something quite unjustified, which is to suppress truths about persons.[2] This puzzling fact animates much of privacy scholarship and has prevented a clear theory of privacy's value from emerging.

Indeed, some have drawn the conclusion that information privacy *is* unjustified, precisely because it hinders access to the truth. For example, some early economic scholarship on privacy argued that information privacy prevents others from making fully-informed and/or efficient choices. According to this line of thinking, explored initially by Richard Posner and Richard Epstein, a job applicant should not be able to keep her checkered past private, because doing so would

[1] Julie Cohen, *What Privacy Is For*, 126 HARV. L. REV. 1904, 1904 (2013) ("No single meme or formulation of privacy's purpose has emerged around which privacy advocacy might coalesce. Pleas to 'balance' the harms of privacy invasion against the asserted gains lack visceral force.").

[2] *See, e.g.,* Robert Post, *Three Concepts of Privacy*, 89 GEO. L.J. 2087, 2088 (2001) (arguing that information privacy is puzzling since it blocks information and information leads to knowledge).

preclude a fairly informed and/or efficient choice by the employer about whether to hire her.[3] Structurally speaking, these views treat persons exclusively as information consumers who are entitled to as much of it as possible, with little or nothing off-limits.[4]

Others take a different tack. They acknowledge that information privacy suppresses the truth, but argue that it is justified nevertheless, because, broadly speaking, persons have interests not just as consumers of the truth in their own right, but as the subject matter of truths known by others.

As attractive as this broad idea may be, it is difficult to extract a unified theory from it. One route to doing so admits that privacy has costs for the truth, but insists that it has sufficiently compensating non-truth benefits, such as, prominently, editorial control over one's life story and so one's public presentation.[5] The idea is that persons are entitled to withhold some information about themselves so that they can exercise control over how others see them, and so, inhabit different personas in their multifarious relationships with others.[6] On this telling, privacy limits access to truth in the interest of what we might call editorial freedom.

This justification is appealing because it purports to identify a species of freedom that privacy protects. However, defending conduct by claiming that one is (editorially) free to do it only

---

[3] Richard Epstein, *Privacy, Property Rights, and Misrepresentation*, 12 GA. L. REV. 455 (1978) (discussing Posner). *See also* Kent Walker, *Where Everybody Knows Your Name: A Pragmatic Look at the Costs of Privacy and the Benefits of Information Exchange*, 2000 STAN. TECH. L. REV. 1, 1 (2000) (not arguing against privacy *per se*, but arguing for a re-balancing of privacy interests given that it "keeps you from enjoying all that society and the market have to offer.").

[4] *See, e.g.*, Walker, *supra* note 3, at 2 (emphasizing the needs of a community to get information about its members).

[5] Legal recognition of the interest in editorial control over one's message has focused on the right against being a host for the messages of others. *See, e.g.*, Hurley v. Irish-American Gay, Lesbian and Bisexual Group of Boston, Inc., 515 U.S. 557 (1995) (parade organizers do not have to admit floats that represent points of view they disagree with); Miami Herald Publishing Co. v. Tornillo, 418 U.S. 241 (1974) (newspaper does not have to include other points of view).

[6] *See, e.g.*, JULIE INNESS, PRIVACY, INTIMACY, AND ISOLATION (1992); Charles Fried, *Privacy [A Moral Analysis]*, in PHILOSOPHICAL DIMENSIONS OF PRIVACY 203 (Ferdinand Schoeman eds., 1984); James Rachels, *Why Privacy is Important*, in PHILOSOPHICAL DIMENSIONS OF PRIVACY 290 (Ferdinand Schoeman eds., 1984); Andrei Marmor, *What Is the Right to Privacy?*, 43 PHIL. & PUB. AFF. 3 (2015).

draws the follow-on question: why do persons have that freedom?[7] I think this follow on question sometimes goes unrecognized, and when it is made explicit, it's hard to see how it will get a good answer because it does not seem that persons have an interest in appearing to others how they wish, but only how they truly are. Indeed, the economic authors I began with make this point after a fashion, since they suppose that information about a person is needed by others (even if they conceive of the need narrowly in terms of economic decisions). To them it seemed that editorial freedom is a kind of deception, and they have a point. To be sure, holding information back is different than outright fabrication, but the point is that desires about how one appears to others do not seem worth indulging absent some reason it would be unfair to appear that way.

A different symptom of this same point is that the editorial rationale is an entitlement to persons to disclose only what they want, but offers no principled stopping point. Sometimes information about persons is needed by others, but the editorial rationale cannot easily say why such a need should be honored, since persons may have just as strong, and perhaps stronger, desire to editorialize on socially important aspects about themselves as the unimportant. Proof of this problematic dynamic is that editorial views often restrict the scope of informational privacy in unprincipled terms, such as to only what is "reasonable."[8] But the question is precisely what privacy is reasonable.

A different approach notices, I think correctly, that information privacy allows persons to escape responsibility for failing to meet various community norms, such as norms about what to do and think.[9] However, this observation by itself does not tell us why information privacy is worth

---

[7] *See, e.g.*, Jed Rubenfeld, *The Right to Privacy*, 102 HARV. L. REV. 737, 750 (1989) (pointing out that invoking autonomy to justify the right to (constitutional, decisional) privacy is a restatement of what privacy protects rather than a justification for it).

[8] *E.g.*, Marmor, *supra* note 6. *See also* HELEN NISSENBAUM, PRIVACY IN CONTEXT 2, 15 (2010) (championing a right to an *appropriate* flow of information based on the need to adjust to different social contexts).

[9] Ruth Gavison, *Privacy and the Limits of the Law*, 89 YALE L.J. 421, 450 (1980).

having, but only prompts a further question of "whether it is appropriate for privacy to permit individuals to escape responsibility[?]."[10]

Once we ask this further question, we are likely to be more puzzled. After all, if we assume that the relevant norms for thinking and acting are sound, such that persons really should think and act that way, then why would it ever be a good idea to let persons avoid being held responsible, such as by censure or punishment, for their failures to meet them? On the other hand, if the relevant norms are unsound, then it seems that it is they – the norms – that should be revised and that privacy can, at best, play a transitional role while that revision is under way.[11] For instance, persons should not be held responsible for their sexuality – because there are no sound norms that dictate what a person's sexuality should be (even if some societies enforce some). But then, it seems that privacy about sexuality is only valuable inasmuch as a "norm" of, say, heterosexuality is wrongly enforced through intolerance or discrimination. If attitudes and laws were to be fully accepting or fully tolerant, then it seems that information privacy about one's sexuality would no longer be necessary and should fade away.[12]

It may seem then that there is a dilemma: that privacy is nothing but an interim, or stop-gap response to mistaken beliefs about what norms persons should be held responsible for meeting, or else, that privacy is nothing but an impediment to holding persons responsible for meeting sound norms. The dilemma is not exhaustive of the options though. It may be that privacy permits persons to escape being held responsible for meeting certain norms, not because those norms are defective or illegitimate, but simply because their very enforcement by others does more harm than good and because persons need space in which to try to meet them, on their own, before being subject to the

[10] *Id.* at 451.

[11] *Id.* at 453 (making a case for privacy on this basis, as a kind of pressure release value for when applicable norms need to be corrected but are slow in changing).

[12] *Id.* at 452.

policing of others. In other words, it may be that privacy stands guard against a kind of over-enforcement of sound norms; of commanding obedience to them too completely or too totally.

I endorse this response to the dilemma, but it is beset with at least two problems. First, many familiar examples of over-enforcing a (sound) norm are contingent, owing only to things like limitations on the efficient deployment of resources or on the chilling effect of deterring a type of conduct with imprecise tools that can be expected to also deter some useful conduct.[13] Privacy, though, purports to be more than that. It purports to be a defense against some norm enforcement on the basis that the enforcement is constitutively excessive and wrongful, and not just inefficient or too costly given other priorities. Second, it may seem that violating privacy cannot be a form of over-enforcing a norm, because in some cases, merely gaining a certain kind of information can violate a person's privacy but gaining information is not a way of enforcing a norm at all.

In this chapter, I will address these two difficulties and others by defending an over-enforcement model of privacy. Because privacy is expansive, I will limit my discussion to one kind of privacy – *mental privacy* – that takes, as its subject matter, a person's mental activity, most prominently, her thoughts, including beliefs, fantasies, fictions, intentions, and feelings. This kind of mental privacy consists of at least two kinds of protections. First, it protects thoughts against being known by others. Consequently, it is representatively violated by obtaining certain kinds of information about thoughts, such as by reading a person's diary or, in science fiction, her mind

[13] For instance, some First Amendment jurisprudence analyzes whether a law that is otherwise justified in restricting speech would scare away too much legal speech. *See* Leslie Kendrick, *Speech, Intent, and the Chilling Effect*, 54 WM. & MARY L. REV. 1633 (2013). A second example: it is sometimes asserted that enforcing civil rights laws against police officers may chill justifiable law enforcement activities. *See* Joanna C. Schwartz, *The Case Against Qualified Immunity*, 93 NOTRE DAME L. REV. 1797, 1811 (2018) (citing this assertion by the Supreme Court but strongly challenging it).

itself,[14] or by identifying her as the author of an anonymous blog post.[15] Second, mental privacy protects thoughts against certain kinds of interactions or uses. It is a violation of mental privacy to punish someone for her thoughts, or to tax her for her thoughts, or, in some cases, to ask for the justification of her thoughts.

One may wonder, why this focus? Why thoughts? My answer is partly practical and partly theoretical. Practically, mental privacy is becoming more vulnerable to violation due to the growing exposure of a person's intellectual products to automatic and algorithmic interpretation as well as to the advances of science in gaining access to thoughts through brain states.[16] Theoretically, I think that mental privacy is likely to provide some guidance about privacy more generally, because the latter often privileges an individual's judgment,[17] and where it does not, such as when it protects perceptions and information about a person's body, the account I offer about thoughts gives a plausible rationale. I return to this claim in the conclusion.

My question then, is: *why is mental privacy justified, viz. why should persons not be held responsible for their thoughts?* Very synoptically, my answer is that persons need space to think by themselves in order to get themselves into shape to think with others. Since having thoughts is how persons think by themselves, and since holding responsible is how others start thinking with them, the latter should not be aimed at the former. Doing so is preemptive or premature, with two consequences. The first is that preempting a person's individual thinking is bound to produce a false impression of who they

[14] For science fiction about mind reading see ALFRED BESTER, THE DEMOLISHED MAN (1951); ROBERT SILVERBERG, DYING INSIDE (1972). *See also*, Jack Yang, *Professor X's Notice and Choice Privacy Dilemma*, THE LEGAL GEEKS, (Aug. 16, 2016), http://thelegalgeeks.com/2016/08/16/professor-xs-notice-and-choice-privacy-dilemma/ (discussing the privacy implications of comic book mind reading powers).

[15] *E.g.*, Alexandra Schwartz, *The "Unmasking" of Elena Ferrante,* THE NEW YORKER (Oct. 3, 2016) (discussing the identification of a popular novelist who had written under a pseudonym).

[16] Kerri Smith, *Brain Decoding: Reading Minds*, 502 NATURE 428 (2013) (outlining some recent results in brain science).

[17] I am thinking here of private property, privacy of religion, and reproductive freedoms (though I share the belief of many scholars that the relationship of reproductive freedoms to privacy is likely to be more oblique than constitutional rhetoric suggests).

are, since it starts thinking about them before they have finished getting themselves ready to be thought about. This would be like evaluating a performance by witnessing a rehearsal for it. To avoid this kind of prejudgment, mental privacy protects a person's thoughts from being known. Second, preempting a person's individual thinking throws them into an activity – thinking together – that they are not ready to participate in, with distorting effects on their reasoning. To avoid this kind of distortion, mental privacy protects a person from being hassled for her thoughts.

This chapter defends this picture across three subsections, as follows. First, I defend some claims about the value and the nature of thinking. I argue that thinking, the continual carrying out of it, is valuable, and that persons carry it out by themselves and with others. Crucially, thinking by oneself is related to thinking with others in roughly the way rehearsals relate to performances, dating relates to marriage, opinion polls relate to election results,[18] drafts relate to published works, or scrimmages relate to games. That is, what a person thinks – her thoughts – relates her to content[19] in a way that simulates, anticipates, or practices the final, official, or public, relationship that she will have to that content in thinking with others, so that she can see and feel what that latter relationship is like before assuming it. This relationship is *being responsible* for some content, which amounts, roughly, to being its spokesperson in a community of thinkers, and hence speaking on its behalf as well as fielding inquiries about it and challenges to it.[20]

Second, I defend a view about what it is to hold someone responsible. Holding someone responsible is, I claim, a stretch of thinking that is initiated by the belief that she is responsible for

[18] Polling is not unlike mind-reading practiced on a collective mind. It reveals, imprecisely, what "the people" think, but not yet what they "say." For the *vox populi*, only the official elections results will suffice.

[19] Content can be a number of things, such as a belief: "global warming is caused by humans," a fiction: "Call me Ishmael," a feeling: "this is fearsome," or an intention "I will go."

[20] *See also*, Cheshire Calhoun, *Standing for Something*, 92 J. PHIL. 235 (1995) (arguing that the concept of standing for something concerns propounding that view in deliberation with the right amount of conviction, where the right amount includes withdrawing it when it is successfully challenged).

some content. For example, a speaker may say that global warming is caused by human beings, and someone may, on the basis of hearing her say that, correctly believe that she is responsible for that content. This belief constitutes the listener holding the speaker responsible for the content about global warming. I should mention that holding responsible often extends beyond just attributing a role to someone, such as when punishment is warranted. My view accommodates this fact, because the belief that someone is spokesperson for some content warrants asking her to speak in defense of that content, which may in turn warrant imposing liability on her if the defense is not satisfactory.

Third, I put the previous two sections together to illustrate the importance of mental privacy. Thinking by oneself is a matter of simulating what it would be like to speak on behalf of some claim, by imagining what can be said for and against it, and then responding appropriately, but all in one's own mind. This process of working things out will be converted from practice to the "real thing" by another's belief that one really is responsible for some content, as this converts simulation to interaction. If this happens when one is not ready, the effect is something like that of an audience walking in while one is rehearsing. In the case of thinking, the result of being caught off guard in this way is twofold. First, one will be seen in a way that misrepresents who one is, in the way that the audience may leave with a misleading impression of one's acting (one was only rehearsing). And second, one's thinking itself may be degraded, in the way that the sudden presence of onlooker may trip one up or cramp one's style by steering one away from trying something out.

I conclude by explaining how the view I defend answers the dilemma sketched above and then by considering how the theory might generalize to bodily privacy.

II. A Theory of Mental Privacy as Protection Against Prematurely Being Held Responsible

A. Thinking: Its Value and Nature

Thinking is a special kind of activity. It has a distinct value and a distinct nature. I will illustrate both in this section. To do so, I will frequently invoke examples that concern only one kind of thinking; namely, thinking about what is true, rather than thinking about how to accomplish a goal or imagining a fictional scenario (though I also try to compactly illustrate the relevance of my argument to these kinds of thinking). I do this for reasons of manageability and also because I suspect that some of my claims, especially that thoughts have the status of experiments or practice, will strike some as most implausible for thoughts like beliefs, which are often characterized flatly as commitments to the truth.

The value of thinking lies in its being the continual exercise or coming to life of a mind and so of a person. In this way, thinking is to persons what homeostasis is to biological creatures, i.e., the active mode of an individual existence. Just as creatures continue living by doing other things, in an organized and related fashion, such as breathing, pumping blood, eating, and so on, so thinking as a whole is sustained by the organization of sub-activities like inquiring, imagining, supposing, deducing, and so on.

This metaphor can be usefully extended one more step. If a creature is deprived of the biological materials necessary for life it dies. Similarly, a person that is starved for "food for thought" will unravel. A mild example is self-quarantining during a pandemic. One may get restless, bored, or a little stir-crazy, without the normal inputs to thinking, most prominently, new

experiences, including interactions with others.[21] If we go further and consider prolonged scientific voyages or solitary confinement, we can notice that without content, in the form of books, conversations, or experiences, the self disintegrates for lack of ideas to chew on.[22] The mind starts to turn on itself, for lack of fresh thinking to do.[23] It has to be kept, as we say, "occupied."[24]

Additional support for this metaphor comes from seeing the ways in which the importance of thinking cannot be wholly reduced to the achievement of some definite, valuable result or product. This is evident when thinking seeks no result at all, such as when one lets one's mind wander. The value of such thinking is in the doing of it, even if it reliably produces all sorts of good things like inspiration for a screenplay, an argument for a paper, or an invention for the patent office.

The importance of thinking itself, the doing of it, remains valuable even when it is aimed at producing a certain result, whether that result is the truth or a convincing fiction. It may seem not so, and that, for instance, thinking that pursues the truth is good only because and to the degree that it will capture its quarry. However, some reflection discredits this idea, because thinking is not just a good or the best means of attaining the truth, but a way or mode of attaining the truth that is, itself, important, and, in fact, the only way or mode in which one can be said to fully "capture" the truth at all.

---

[21] *See What We've Learned*, N.Y. TIMES MAG. (May 19, 2020), https://www.nytimes.com/interactive/2020/05/19/magazine/covid-quarantine-dust.html (a collection of short essays about COVID-19 quarantine).

[22] SUE HALPERN, MIGRATIONS TO SOLITUDE 49 (1992) (researcher who isolated herself to study social connection reported numerous problems, such as "dull numbness" and "loneliness"); SEANA VALENTINE SHIFFRIN, SPEECH MATTERS 90-91 (discussing solitary confinement); Rachel Aviv, *How Albert Woodfox Survived Solitary,* NEW YORKER, Jan. 8, 2017 (discussing solitary confinement).

[23] HALPERN, *supra* note 22, at 53 ("solitary confinement is like having hell all to yourself").

[24] A caveat: I am not trying to argue that one must be, at every moment, thinking. Sometimes, we want to dampen our thinking or to think about nothing at all, whether by meditation, insipid television, sleeping, or alcohol. These however, are temporary breaks from thinking that one chooses, and which do not preclude one from resuming it again.

Take reaching a mountaintop for a comparison. One can get there by helicopter or by walking, but the latter is a way of arriving that realizes or captures a distinct (not necessarily superior in this case, just distinct) value to being at the top. A mark of this fact is that arriving at the top of a mountain by walking is a *summitting* in a way that arriving by helicopter is not. Thinking is also a way of arriving at some result that makes the arrival distinctly valuable as compared to other ways of doing so. We might term the arrival at a particular conclusion about what is true that is brought about by thinking a *realization*, or an *understanding*.

The comparison is imperfect in two ways. The first is that a life without any summitting, or any mountain climbing for that matter, can be as good as any, but a life without any realizations or understandings, and hence no thinking, is seriously defective, even if it somehow achieves the truth. I say "somehow" because, and this is the second way the analogy is imperfect, reaching the top of a mountain is something that climbing makes happen, but possessing the truth is not something that thinking makes happen. The reason is that possessing the truth is not something that consists of a thing, a body, being (passively) in a particular location, but involves a mind relating to something through activity. That is, possessing the truth is, itself, a bit of thinking held stationary, or in equilibrium, ready to be deployed.[25]

Examples testify to these claims abouts thinking. Consider curiosity and creativity. These are good traits to have and good as attractions or attunements toward the activity of exploration. Curiosity is good because it is a general susceptibility or a sensitivity to the value of the truth and the pleasures of searching for it.[26] Someone who is curious does not simply want the truth, but rather is easily moved to start thinking or wondering about it, and this is valuable in itself in a way that is not

[25] Matthew Boyle, *Two Kinds of Self-Knowledge*, 78 PHIL. & PHENOMENOLOGICAL RES. 133 (2009) (discussing belief as an activity); Pamela Hieronymi, *Believing at Will*, 35 CANADIAN J. PHIL. 149, 174 (2009) (making a similar point).

[26] *See* Elias Baumgarten, *Curiosity as Moral Virtue*, 15 INT'L J. APPLIED PHIL. 169 (2001).

undermined by a failure to actually hit on the truth. As teachers know, the formulation of a question is an important learning moment, even if the answer escapes the student, because it evidences a desire for truth, by doing some thinking, in the formulation of the question, to try and get it.[27]

Creativity is also an affinity for exploration, through the exercise of one's imagination.[28] Creativity does not seek to discover what is true but to see what is possible and, in some cases, possibly meaningful, beautiful, useful, and so on. The pursuit of this goal reveals or exhibits potentiality; of one's mind as revealed in the medium or structure being manipulated. For instance, a child playing with blocks attempts to see what can be done with them and finds that they can be a house, which exhibits the potential of the blocks and of her own mind acting through them. Poetry attempts to do something, where the bits to be manipulated are words.[29] Of course, which "bits" are in play may itself be a matter of creative choice. Here too, as with curiosity, a failure to create something non-derivative, beautiful, or useful is not fatal to the value of the endeavor. The aspiration and effort themselves are valuable as attempts to deploy and realize potentiality of a rational character, where rational is not meant to concern truth or inference, but of significance or meaning.

---

[27] Not all questions are like this. Some questions are asked to see if the other person knows the answer (as on a quiz or test), but some are "genuine" or sincere in being the desire for the truth, formulated in terms that one thinks are productive to finding it, such as when a scholar tries to formulate a research question that needs to be addressed by a field. On types of questions and their role in learning, see ANN SHARP & LAURANCE SPLITTER, TEACHING FOR BETTER THINKING (1995).

[28] Note too that imagination plays a crucial role in many kinds of thinking – we may imagine possibilities to test ideas, and we imagine courses of action to secure practical results. However, imagination is not limited to such auxiliary uses. It can be used "purely," to see what it brings. Marc Blitz, *The Freedom of 3d Thought: the First Amendment in Virtual Reality*, 30 CARDOZO L. REV. 1141, 1172-3 (2008) (discussing the use of imagination and its relationship to privacy), ANTHONY STORR, SOLITUDE: A RETURN TO THE SELF (1988) (discussing the way privacy can aid creativity); Barbara Herman, *the Morality of Everyday Life*, 74 PROC. AND ADDRESSES OF THE AM. PHILOSOPHICAL ASS'N 29 (2000) (discussing creativity, its connection to self-absorption, and the moral risks it may consequently pose).

[29] TIMOTHY MACKLEM, INDEPENDENCE OF MIND (2008) (focusing on languages, natural and otherwise, as providing a set of constituents whose potentiality can be exhibited through their creative deployment).

A second example is a pair, that of being told what is true and being forced to believe what is true. Consider first being told something. Clearly, one can justifiably believe what one is told.[30] However, being told something, or even being told something along with the reasons that support it, is no substitute for arriving at the truth by following out the reasons for it. Knowing by being told is, literally, second-hand thinking – one is dependent on the thinking of another. Teachers tell students truths, but they are often cagey about it, because it risks encroaching on the student's chance to work it through.[31] Moreover, when a teacher does tell a student something true outright, the hope is that the student will adopt it, skeptically, until such time as they can certify it for themselves, or at least, to take it with a grain of salt.[32] For these reasons, there is an ethics of telling.[33] Telling is unobjectionable and most at home when one person must relay a firsthand experience of something the other could not have experienced (such as an eyewitness account of some event), but it can be patronizing, such as when one tells someone something they can be expected to know for themselves, or something that they can be expected to take issue with.[34]

If being told is a touchy, but sometimes-justified form of voluntarily out-sourcing one's thinking to another, then paternalism and mind control are unilateral and forbidden forms of take-over. The paternalist "substitutes" his judgment for the victim's, depriving her of a chance to think, and this is wrong even when the victim can be expected to make a mistake according to some metric. Similarly, being compelled to recite claims (even if true), or to suffer a science-fictional kind

---

[30] *See* Tyler Burge, *Content Preservation*, 102 PHIL. REV. 457 (1993) (discussing the basis for believing what others say).

[31] Or think about working on a crossword puzzle, and receiving the answer to a particular clue, unsolicited, from someone sitting nearby. Here, the problem is chiefly being robbed of the fun, or entertainment value of thinking, so it is illustrative, but parasitic on the central claim that thinking, even when it's not for fun, is something that one should be able to do.

[32] Teachers often ask students to put a point in their own words, as a minimal but efficacious first step toward building their own understanding of it. Many tests do not ask for revelatory new arguments from students, but rather, ask only that students speak out ideas from the course in a way that bespeaks their own organization and internalization of the material. *See* George Tsai, *Rational Persuasion as Paternalism*, 42 PHIL. & PUB. AFF. 78, 99 (2014).

[33] *Id.* at 91-93, 103.

[34] *Id.*

of brainwashing (even if one is brainwashed to believe the truth) is wrongful because it infringes on the victim's thinking by taking it over, or bending it subconsciously toward a particular conclusion.[35]

With some exceptions, I've been illustrating the value of thinking with examples that concern thinking about what is true, but the former is as expansive as there are kinds of thinking, and there are many. One can think by dreaming, or by planning to pull something off, by identifying as something or with someone or something, by feeling and emoting, or by being in a mood. All of these are ways of coming to grips with something, where this is accomplished through the restructuring of oneself or one's viewpoint, at various levels of consciousness or explicitness and with differing connections to other thoughts and actions. A mood is an inchoate readiness to be moved, such as to feel sadness, and sadness may lead to crying or tiredness. Identifying with or as something is more explicit in that it involves taking on an aspect, or feature, out of recognition of its fittingness, where the aspect often has or is evolving, a socially recognized meaning, e.g., identifying as Catholic or gay. Dreaming may be phenomenally quite impressive, but the working through that it represents is not explicit in the sense that a dream does not announce any conclusion or idea, but is fodder for interpretation. Sadness too must be interpreted (*why* am I sad?), but there is a core to sadness that announces its meaning – something is wrong. Planning, or designing a machine to do something, is like trying to find the truth, in being the mobilization of thought *for* a goal. I don't think one feels sad to reach a conclusion, but it does help toward that end, among others.

Many recognizably important activities mobilize these and other ways of thinking in patterns and packages that give them added potency. Take grieving. It is a process by which one will be in moods, feel things, concentrate on various thoughts, alter one's identity (e.g., "widow"), where these things, ordinarily, serve, to help one understand the importance of the deceased in one's life. I'm not

[35] Seana Valentine Shiffrin, *What is Really Wrong with Compelled Association?*, 99 NW. U. L. REV. 839, 852-60 (2005) (discussing compelled speech).

saying one grieves in order to learn something – that would be an oddly intellectual and self-focused way of viewing it – but rather that one learns something in the process of trying to arrange oneself for the future, to incorporate and "keep alive" the memory of the one has passed. It's a way of preparing an edited and vibrant simulation of the deceased person's self, that one can keep "running" in one's own mind, after they're gone. This is just one of many activities by which we grow as persons.

At this point, I take myself to have vindicated the illustrative metaphor I started with, to some extent, and so shown that thinking is broadly valuable, as it is the signature activity by which a person exists and grows. I now turn to the nature of thinking, which is to say, its fundamental structure. My focus will be on how it is conducted and the way in which it constitutes a person.

A good way into my main idea is by considering practicing a sport like basketball. Two things are relevant. The first is that practicing is a way of gaining knowledge, specifically, practical knowledge of how to do something well. One practices something by trying, as much as possible, to do the thing one is practicing to do. So, one practices shooting well by trying to shoot well, even if one, in fact, shoots poorly and so misses. Practicing is a way of learning because one can learn from a missed shot just as much as from a made shot. Indeed, one often learns more from a failure than a success, because it signals how to recalibrate. This latter idea, to its credit or discredit depending on your perspective, can be found in an uncountable number of biographies and self-help books. I think Einstein put it well when he said (allegedly) that "failure is success in progress."[36] We learn from failure.

[36] This statement is attributed to Einstein by a number of sources, but without citation. *See, e.g.*, Meredith Young, *The Utility of Failure: A Taxonomy for Research and Scholarship*, 8 PERSPECTIVES ON MEDICAL EDUCATION 365 (2019).

Second, there is a distinction between practicing an activity and "really" doing it.[37] In the case of sports, this distinction is largely conventional in that what makes something a real game is its designation as real by an organized system of play, say, the NBA. What is interesting and suggestive though is that the realness of sports – which games count – is partly a matter of how accessible it is made to others, which is in turn a function of how many others are interested in the result. Hence, sports, conventional as they are, achieves the "realness" of its games by tapping into the significance that observation has for persons, seeking out and also satisfying the desires of others to see what happens in a game, which is reinforced by the keeping of records; namely, "official" stats. Sports are a derivative example of a theme, which is that for persons *publicity is realness* because of the significance that others have for us, because publicity is vulnerability to being known by others, and because something exists for a particular mind or thinker by its being known by it.[38]

Thinking is like practicing in both of these regards, but, in a sense, even more so, because unlike basketball which is a derivative use of publicity, thinking by oneself is constitutively a practice for social thinking, which is not just a popular pastime that some enjoy, but has a significance for persons that makes it true that public thinking is "real" thinking that "counts." Shooting a shot on TV, for the NBA, is a "real" shot as opposed to a practice shot, not in any intrinsic way, but only through the deliberate marshaling of attention and interest, the creation of a schedule, a league, and so on, but thinking that is known by others is real, inherently.

---

[37] Before playing a game, one may take a practice shot, and this does not count, which is to say, it's not real, for the purposes of the game.

[38] HANNAH ARENDT, THE HUMAN CONDITION 45-46 (1959) ("For us [humans], appearance...constitutes reality. Compared with the reality which comes from being seen and heard, even the greatest forces of intimate life -- the passions of the heart, the thoughts of the mind, the delights of the senses -- lead an uncertain, shadowy kind of existence unless and until they are transformed, deprivatized and deindividualized, as it were, into a shape to fit them for public appearance ....Each time we talk about things that can be experienced only in privacy or intimacy, we bring them out into a sphere where they will assume a kind of reality which, their intensity notwithstanding, they never could have had before.").

The reason is that thinking by oneself is not a solitary activity that one can make known to others, but rather, the solitary sustainment of an inherently witnessed, dialogical activity – like promising oneself, playing chess with oneself, or playing catch with oneself. In solitary thinking, one is playing "both sides of the board" in that one becomes responsible for content, if only to ourselves, in the sense that we believe something so long as we can answer, to our own satisfaction, the challenges we can create for it. We often say that considering something is a matter of seeing "what can be said for it" or "on its behalf." When a person finds out what we are thinking, they immediately take up one of the two perspectives that we were occupying, and transform our activity with ourself into an activity with another, with exactly the "board state" we were facing before they came along. Practice becomes real.

If we later see that we missed a serious challenge or abandoned some thought when more could be said for it, we have made a mistake, but we can learn from failing, and this too is a holding ourselves responsible. Failed thinking is thus rational success in progress. We can see this point in greater detail by further comparing interpersonal thinking with intrapersonal thinking.

The importance of misunderstandings in reaching understanding is pointed out, famously, by J.S. Mill when he notes that understanding the truth, really fathoming it, is furthered by confronting someone who sincerely holds the opposite view.[39] Many intellectual practices are modelled on this kind of confrontation. Free speech at the societal level is valuable in that it tries to ensure that dissident views can challenge and enrich orthodoxy, preventing the latter from hardening into dogma. Switch-side debating, in which participants take turns advocating both sides of a resolution (for and against it), seeks to produce understanding out of different perspectives, as do

[39] JOHN STUART MILL, ON LIBERTY 34 (New York: Penguin Books 1985).

adversarial legal systems, in the pursuit of justice.[40] These forms of collective thinking are valuable because not only may one participant confront another with an opposing view, but she may also have to inhabit the opposing view, and so come to grips with it "from the inside." A similar practice is the appointment of "devil's advocates" to argue against the canonization of saints, to ensure that the final decision was robustly supported.[41]

Importantly, the foregoing examples involve multiple persons holding different views, with different levels of sincerity, at a single time. I wish to stress a slightly different but related point, which is that the same benefits to understanding redound to individual persons who sincerely hold different viewpoints across different times. For example, switch side debating is an imperfect substitute for being actually mistaken and then changing one's mind, because in being mistaken, one is "taken in" by the false view, and so knows its attractions more intimately than by just mouthing it. One inhabits a mistaken view most deeply by actually believing it, and not just by trying to see what can be said for it.

I'm not arguing that persons should try to succumb to false views. Rather, my point is that the mere holding of a false view is not, itself, at odds with understanding the truth, but is, given the never-ending nature of thinking and the enriching effect of false perspectives, a "success in progress." Just believing something false, due to being an ongoing bit of thinking, keeps open the chance to see what is true, and prepares for that realization to be one that is not idle, but backed up

[40] *See* Casey Harrigan, *Against Dogmatism: A Continued Defense of Switch-Side Debate*, 29 CONTEMP. ARGUMENTATION & DEBATE 37 (2008).

[41] *Id.* at 49. An interesting further example is the Jewish practice, *Beit-Din*, of not permitting unanimity in the imposition of the death penalty, since it would be evidence that the requisite clash of arguments had not occurred. *See* Edna Erez, *Thou Shalt Not Execute, Hebrew Law Perspective on Capital Punishment*, 19 CRIMINOLOGY 25 (1981).

and enriched by one's struggle to secure it. We sometimes have to learn "the hard way" (by trying and failing), and when we learn this way, we get what is often called wisdom.[42]

A second example of the value of mistakes comes from considering works such as art and writing, which are products of the rational activity of their author or maker.[43] Writing something down is a provisional kind of commitment to it. By putting an idea on paper, one captures it, but not so that one can wholeheartedly endorse it, but only to get a better look at it, like holding a specimen under a magnifying glass to see whether it's the right one.[44] If, as is frequently the case, one is not satisfied with the idea as written, one has not completely failed to make one's idea clear or persuasive. Instead, the draft is itself progress, as it will *live on* in the final product.[45] Moreover, in art and in writing, it is the blank page or canvas that is most intimidating, because one has no access point or foothold into the work that it will become. Once one makes the first mark, or writes the introduction for the first time, one has something to think about; a target for rational attention. Even if it's a mistake and one erases the mark or rewrites the introduction, the mistake anchors subsequent engagement in a productive way.

Now let me turn to showing that thinking by oneself is not "real" and so does not constitute who one is, but only an anticipation or practice version of who one is. At any given time,

---

[42] *See* Todd D. Rakoff, *The Implied Terms of Contracts: Of 'Default Rules' and 'Situation-Sense', in* GOOD FAITH AND FAULT IN CONTRACT LAW 191 (Jack Beatson & Daniel Friedmann eds., 1995) (arguing that persons obtain wisdom by playing different relational roles, such as parent, friend, employer, trustee, etc.).

[43] *See* Keith Oatley & Maja Djikic, *Writing as Thinking*, 12 REV. GEN. PSYCHOL. 9 (2008); Keith Oatley, *Thinking Deeply in Reading and Writing, in* THE EDGE OF THE PRECIPICE 175, 180 (Paul Socken ed., 2013) (quoting Richard Feynman as responding to an interviewer's statement that his notebook was a record of his work, he responded "it's not a record, not really. It's *working*. You have to work on paper and this is the paper.").

[44] *E.g.*, Roger Traynor, *Some Open Questions on the Work of State Appellate Courts*, 24 U. CHI. L. REV. 211 (1957) ("I have not found a better test for the solution of a case than its articulation in writing, which is thinking at its hardest. A judge...often discovers that his tentative views will not jell in the writing. He wrestled with the devil more than once to set forth a sound opinion that will be sufficient unto more than the day.").

[45] As I will argue in the next paragraph, beliefs are like drafts in being provisional. One may assume something to see what it entails, and certainly beliefs are not like that, since one does not adopt beliefs, on purpose, to see where they will take one. Still, beliefs are, I claim, a provisional mode of being connected to the truth in being inherently revisable.

a person's identity is largely a matter of what they think, i.e., their thoughts.[46] However, we have to be careful how we understand the way a person is made up of their thoughts, because we may be tempted toward a static picture in which thoughts are things and persons are the vessels in which they reside, when in reality thoughts are the equilibria shape that the person's thinking happens to take. On this latter picture, a person is not just a container for thoughts, antecedently existing and being at first filled with rough thoughts and then, with any luck, more refined ones, but rather a single rational perspective that progressively comes into existence through, or in, the progression from the rough thoughts to the more "thought out" ones, in the way that a novel takes shape through the continued replacement of bad drafts with better ones.[47]

Notice that practicing has no end point, in itself. There is no point at which the goal of practicing is definitively obtained. Rather, what happens is that one makes a decision that one is *ready* or that one has *practiced enough*. This is not a matter of practicing coming to its own end, in the way that house-building ends itself when the last nail is hammered in. House-building brings itself to an end because its constituted by a striving toward a goal that, in principle, can be reached. Instead, ending one's practice is a decision that one is ready and willing to perform one's skill for real, with whatever consequences that will entail.[48]

To elaborate: when an activity comes to an end and has a product, the product, we say, is a finished product. Manufacturing a car is an activity that comes to an end and has a product, and so

---

[46] Seana Valentine Shiffrin, *A Thinker-Based Approach to Freedom of Speech*, 27 CONST. COMMENT. 283, 291 (2011); JANNA MALAMUD SMITH, PRIVATE MATTERS: IN DEFENSE OF THE PERSONAL LIFE (1997) ("Together with their particular biology and historical moment, people are ultimately made individual by those parts of experience that are difficult to communicate, the images sorted through as sleep comes, the compose of coherent memories, fragmented pictures, habits of love, powerful feelings, and perceptions that, like so many leaves raked into a pile, arrange themselves in relationships that cannot be replicated or completely described.").

[47] HERMAN HESSE, STEPPENWOLF 62 (trans. Joseph Mileck & Horst Frenz, 1963) ("…man is not yet a finished creation but rather a challenge of the spirit; a distant possibility dreaded as much as it is desired…").

[48] Again, in sports, performing for real might lead to embarrassment or even humiliation, but one's desire to be a competent sports player, though important, is of a different kind than one's interest in having the position one wishes to occupy in thinking for real with others.

the output of this process – the car – is a finished product. But the product of an activity that evolves but does not come to an end is not a finished one, but is instead a work in progress, or under development. Thoughts are this way. As the product of thinking, they are always in progress, and hence the person who thinks them is, as a consequence, always in the middle of becoming who she is. The mind is a laboratory, a studio, or a workshop, and in it, persons are always making themselves, never made.

It may seem that this fact has the radical and unattractive consequence that we never really interact with persons, since they are never there, in their entirety, to be interacted with. However, this conclusion does not follow because persons do not think just by harboring thoughts about what is true, but by deciding when and how to adopt the content of those thoughts as aspects of oneself, with the latter being an act of assuming responsibility for the content, as oneself, in one's entirety. In this way, thoughts are, as the products of thinking, just like books or works of art. They have no point at which they are complete, and yet, one will make (an often agonizing) decision to force a kind of completion on them, by putting them out into the world in one's own name, so that others can in turn make them the subject of *their* thinking.[49] In this way, a work (of the mind, e.g., books, performances, speeches, sculptures) is a caged infinity – a whole way of thinking that has been forced into an uneasy truce with a finite form so that others have a door through which to enter it.[50]

The interaction of these two aspects of thinking – having thoughts about what is true, and deciding to adopt the content of those thoughts as parts of oneself – constitutes persons as both always making themselves in thought, and also, of having partially made themselves through their

[49] K.E. GOVER, ART AND AUTHORITY (2018) (surveying the philosophical debate about when a piece of art is finished); Ann Landi, *When Is Artwork Finished?,* ARTNEWS February 24, 2014 (interviewing artists explaining the various ways that they decide when a piece is done).

[50] Alexander Nehamas, *An Essay on Beauty and Judgment,* THE THREEPENNY REVIEW 2000 (explaining that good art is never fully understood, but provokes further thinking the more it is appreciated).

choices of which thoughts to adopt. In this arrangement, one's mind is the site of perpetual experimentation, the results of which one decides, from time to time, to bring out into the interpersonal world constituted by relationships with other persons. Thus, we do interact with persons in their entirety. We do so by interacting with the parts of them that they have thoroughly "thought out," i.e., those thoughts that they have brought out of their mind for us to interact with.

A simple example is ordinary, non-speech action. Imagine that one is thinking about whether to cut in line at the grocery store. At some point, such thinking may progress to an intention to cut, and then to cutting.[51] The cutting, I claim, is the adoption of the content, roughly, that cutting is what one should do here and now, in form that is not a thought. This kind of adoption makes it no longer provisional or experimental. One may be held responsible for cutting, as a concretized, settled form of one's thinking about what to do.[52]

Two points are needed. The first is that adopting content in this way is not forever. One may at some point, rightly say "I'm not that guy anymore" such that the content of past action is no longer a part of oneself.[53] The second is that making content a part of oneself, for as long as that lasts, should not be understood as finishing, ending, or even pausing the thinking that led to it. One can still think about whether cutting in line was the right thing to do, even after cutting, and indeed, doing so may be important, as part of apologizing is rethinking what one did, seeing it as mistaken, and then trying to adopt the content of one's re-thinking – that one was mistaken – through the act

[51] I am here skipping over a range of interesting issues regarding intentions. They present special issues because they are halfway between mind and world. They are creatures of thought, and so revisable, and so somewhat experimental, which is the basis of their instability for others, as compared to promises. However, they are also commitments to act, and for that reason resist being revised in various ways, on pain of sabotaging their mission of readying action.

[52] Jenna Donohue, A Deliberative Conception of Complicity (forthcoming) (unpublished Ph.D. dissertation, University of California, Los Angeles) (on file with author) (arguing that actions communicate content about what is acceptable to do).

[53] *E.g.*, George Kateb, *On Being Watched and Known*, 68 SOC. RES. 269, 279 (2000) ("The second consequence is that a detailed record follows a person through life, growing old with him or her, yet not losing memory as the person does. A person will not be able to start life over again, free of some of time's filthy load.").

of apologizing.[54] One cannot go back in time to change one's wrongdoing, but one can always change one's thinking, and if one does so, and adopts the change through the act of apologizing, then one goes a long way toward repairing the moral damage.[55]

The role of action in adopting thought contents is even more strikingly instantiated in the special case where action takes the form of speaking sincerely.[56] Sincerely speaking the content of a thought does not just provide a hearer information about the thought, in the way that a picture provides information about the thing pictured. Rather, sincerity is a mode of relating to content, by which one puts oneself fully behind it, and incorporates it more strongly into one's identity. Speaking is how one *takes* a position, in the sense of grasping it and holding it openly and actively, which is a solicitation to others to take account of it, which, if they do so, makes it one's real position, just in virtue of being known to be such among thinkers. By playing against others, one plays a game rather than a bit of solitaire, and by thinking with others, one is a spokesperson rather than just one in training.

Evidence for this view of things is that providing information about a thought does not have the same interpersonal significance as a sincere assertion of its content, in the way that giving information about an intention does not have the same interpersonal significance as a promise. Saying "I believe the bus comes at 2," is different from saying "the bus comes at 2." The first (sincerely) reports the existence of the belief that the bus comes at 2, but the second sincerely asserts that belief's content. The former is a hedge, and the second is a commitment. Similarly, prefacing a

---

[54] *See* Jeffrey Helmreich, More than Words: Stances as an Alternative Model for Apology, Forgiveness and Similar Speech Acts (2013) (unpublished Ph.D. dissertation, University of California, Los Angeles) (on file with University of California eScholarship) (available at, https://escholarship.org/uc/item/3z03x8qv).

[55] *See* Linda Radzik, *Tort Processes and Relational Repair, in* PHILOSOPHICAL FOUNDATIONS OF THE LAW OF TORTS 233-235 (John Oberdiek ed., 2014) (apologies are aimed at repairing a wrong itself and not its collateral consequences).

[56] I'm relying heavily here on the convincing work of Richard Moran about speech, across several publications. See Richard Moran, *Problems of Sincerity*, 105 PROC. ARISTOTELIAN SOC'Y 341 (2005); Richard Moran, *Getting Told and Being Believed*, 5 PHILOSOPHERS' IMPRINT 1 (2005); RICHARD MORAN, THE EXCHANGE OF WORDS (2018).

statement with "I'm thinking out loud here" is a way of hedging, by indicating that one is only pulling the curtain up on one's mind so that the thoughts are revealed rather than being sincerely projected outward.

The kind of commitment worked by speech is like that of non-communicative action in being time-limited and allowing rethinking. One's words from years ago may be stale, and even if they are recent, they are not binding. One can rethink content that one acts on, and one can say something different after making a statement.[57] The point though is that in saying something, it becomes, for a time, a part of one's biography or one's record, and if one says something different, this is a change to one's full judgment or position and not the relinquishment of something tentative. For example, a politician is not a flip flopper because she changes her mind several times before "coming out" in favor of something, but only when she changes her *public* position many times, or changes it without a clear reason, as this invites the charge that she does not think things through before speaking, or that she takes positions cynically or opportunistically. Hillary Clinton favored the Iraq war in 2002 (and voted in favor of the use force resolution) and though she has distanced herself from that position, it represents her judgment in a heightened way that mere thoughts about the war do not.[58] A public judgment is judgment for real.

To summarize: I have presented the following picture of thinking. Thinking is a valuable, continuous, and incrementally progressing activity of finding one's way to the truth and then maintaining oneself there. It works by a process of accumulation, or as I have said, by conserving one's past judgments, inferences, etc., in one's current thinking, lending it a kind of richness and

[57] In some cases there may be residual duties to inform those who have relied on previous speech, but the duty to inform is different than the duty to maintain a position.

[58] Scott Beauchamp, *Why Clinton's Iraq Apology Still Isn't Enough*, THE ATLANTIC (Sept. 8 2016), https://www.theatlantic.com/international/archive/2016/09/clinton-iraq-bush-war-hussein-wmd-senate/499160/ (Bernie Sanders talks about Clinton's vote as shedding important light on her judgment)

depth that is often labeled as wisdom. Importantly, the product of thinking is thoughts, which in turn make up a person. Hence, persons become more real, more truly themselves, as they do more thinking, with the adoption of a particular thought's content, by taking responsibility for it, a way of making it a full part of oneself.

Let me finish this section by forestalling a misinterpretation of my argument. Some argue that persons are not responsible for at least some of their thoughts, such as beliefs.[59] This is not my position. Persons are, in the main, responsible for the content of their thoughts.[60] Hence, by taking responsibility for the content of a thought, a person does not create a relationship to that content that was not already there. Rather, taking responsibility solicits others to take notice of the relationship one has to some content, where their taking notice makes one's responsibility for the content more urgent or significant by having interpersonal consequences. Their taking notice makes one's affiliation to the content real.

B. Holding Responsible: Thinking With Others

In this section, I defend a view about the nature of holding-responsible – about what it is. Some think that holding responsible is a matter of doing something to someone that is appropriate, given what they are responsible for. So, if someone is responsible for something meritorious, praise may be the right way of holding them responsible, but if they are responsible for wrongdoing or lawbreaking, then blame or punishment is appropriate. I will offer a different view according to which holding responsible is the recognition of a person as a participant in a community of thinkers and thus that their viewpoint (defined by the content they are responsible for) must be taken

[59] *E.g.*, Douglas Husak, *Does Criminal Liability Require an Act?, in* PHILOSOPHY OF CRIMINAL LAW: SELECTED ESSAYS 48 (Douglas Husak ed., 2010).

[60] One argument among several for this claim is a *reductio*. If persons were not responsible for their beliefs, how could they be responsible for much of anything? Pamela Hieronymi has argued for this conclusion across several publications. *See Pamela Hieronymi, Controlling Attitudes*, 87 PAC. PHILOSOPHICAL Q. 45 (2006); Pamela Hieronymi, *Responsibility for Believing*, 161 SYNTHESE 357 (2008). *See also* Robert Adams, *Involuntary Sins*, 94 PHIL. REV. 3 (1985).

account of – addressed and reckoned with, though not necessarily accepted – in the course of working out a mutual understanding. In being this kind of recognition, it is also, itself, the beginning part of that process of working out a mutual understanding.[61] By recognizing that someone is responsible for an idea, one puts that idea, as that person's, on the radar of the community as one to be incorporated into group thinking, such that criticisms of it should be directed to her as well as requests about its merits.

Understood this way, holding persons responsible is the gateway to collective thinking. It is the way in which one person (or persons) takes an interest or takes account of the conduct or thinking of another and makes it her "business." When it is done properly, practices of holding responsible are ways of sharing responsibility for group decisions. They help to constitute a fair community and to effectively pool the powers of individuals that comprise it. For example, communities of scholarship create better thinking by holding participants responsible for what they propose, and a commonwealth pools the resources of its members by holding persons responsible for what they own and buy and sell. Rule of law societies create a fairer form of social life by making each citizen responsible to each other for how they act: no one is above the law.

Admittedly, these are highly abstract claims, and fully defending them would overwhelm the goals of this chapter, nonetheless, they can be made plausible with some more concrete claims and examples. Specifically, I will argue that holding a person responsible is a stretch of thinking that always begins with, and may sometimes be completed by, the warranted belief that someone is responsible for something. This judgment is, as we have seen, the judgment that a person can respond to interlocution regarding the thing they are being held responsible for and thus that it can be taken account of in the creation of a mutual understanding, on the basis that it is not simply a

[61] *See* Helmreich, *supra* note 54 (explaining the idea of a stance, which is a commitment to do something that also does something).

"dead" bit of content, but has a "live" sponsor or spokesperson who makes it responsive to deliberation.

Consider an ordinary interpersonal case. One is standing in line and watches as someone cuts. Here, one gains *prima facie* evidence that the person is responsible for cutting, and since cutting ordinarily constitutes a wrong or a slight, one has the subject matter and warrant for asking the question "What do you have to say for yourself (regarding your cutting)?"[62] The cutter may deny one's version of the facts ("It only looked like I cut, this line was open.") or concede one's versions of the facts, but deny responsibility through an excuse ("I didn't see this was the line, the sign is misleading"), or, concede that he is responsible for cutting, but then offer a justification ("I was told to go to the front because these supplies have to get out ASAP"). In the latter case, one has to hear the justification and evaluate it, which may require further discussion with the cutter. If one does not credit the offered justification, then one may be warranted in asking for an apology. The point is that one justifiably believes that the cutter is responsible for cutting, and truth of that belief then makes it reasonable to ask the cutter to defend the cutting (here a defense is asked for because cutting is *prima facie* wrong), and the failure of the cutter's justification may in turn warrant the request for an apology, the giving of which may create a mutual understanding. If that fails, then steps have to be taken that are not steps in a discussion, but are more akin to imposing liability, such as, in this case, blaming the cutter and/or asking for store employees to get involved.

Institutional practices of holding responsible follow this basic pattern. A good example is criminal punishment. By constitutional command, a federal criminal charge usually begins with a grand jury finding that there are grounds to believe that someone did something illegal.[63] This is a

[62] In a real social situation involving adults, such a phrase would be out place, given that it's usually used to scold children. Perhaps one might say "Hey, there's a line here."

[63] U.S. CONST. amend. V ("No person shall be held to answer for a capital, or otherwise infamous crime, unless on a presentment or indictment of a Grand Jury ....").

finding that there is a sufficiently strong *prima facie* case of lawbreaking, based on evidence that the grand jury is entitled to consider. This *prima facie* finding of responsibility then warrants a deeper inquiry, which involves an invitation to the defendant to explain themselves, usually through an attorney's orchestration of their defense, which may take the form of a denial of the facts, or an excuse, or a justification. The court listens, and if the defendant's answer is unsatisfactory, it responds in turns with penalties, but also, crucially, with an explanation of its own (a reasoned opinion), that completes an interlocution with the defendant and explains the reasoning behind the punishment.[64]

My point with these examples is that an episode of holding responsible progresses toward a mutual understanding, starting from available, but hardly dispositive evidence about what was done by whom, and its appropriateness according to various standards. In the interpersonal example, it is only because one has (and is entitled to have) some evidence that someone is responsible for something *wrong* (cutting) that one needs to dig deeper, by asking them to explain themselves,[65] and it is only, in turn, when and because the explanation is defective is one warranted in requesting an apology. This brings me to a different claim, which is that not all the stages of holding responsible are always necessary. That is, in some cases, when no further stages are warranted, an episode of holding responsible may start *and end*, as a complete instance of holding responsible, with nothing more than the judgment that someone is responsible for something, because such a belief is, under

[64] Interestingly, in the interpersonal case, one demands an apology which one may then accept. But punishment is something that is delivered to the defendant as something that *he*, not the victim, may accept. *See also*, Andreas Schedler, *Conceptualizing Accountability, in* THE SELF RESTRAINING STATE: POWER AND ACCOUNTABILITY IN NEW DEMOCRACIES 13 (Andreas Schedler et al. eds., 1999) (holding someone responsible is an exercise in offering reasons).

[65] This itself is a microcosmic instance of the function of privacy. One needs evidence that something wrong has happened before one can ask for or demand an explanation, because the latter requests access to a person's reasoning, itself, which is protected by mental privacy. An institutional example is the search warrant. The government needs some evidence that one is involved in something illegal before they can search one's house, because a house, like one's mind, is private and so can't be accessed absent some evidence that one will find something important.

the circumstances, a sufficiently robust mutual understanding, i.e., how to allocate responsibility to persons for actions.

This form of holding responsible – a belief that someone is responsible for some content – I will say, is the creation of a record (or an account). In interpersonal situations, records may take a mental form. For example, two Alcoholics Anonymous members may hold each other responsible by periodically checking in about each other's sobriety. Similarly, two colleagues may hold each other responsible for writing a certain amount by checking in at the end of the week, where a failure to write need not warrant an explanation or even blame. Moreover, persons may publicize New Year's resolutions so that friends and family hold them responsible for meeting them, by knowing if they succeed or not, whether or not this occasions anything further, such as a demand for justification (which it usually will not).[66]

Institutional examples can be found as well. Congress may launch an inquiry, audit, or investigation into what a particular agency did in response to some event of public import, and even if it does not find any improper conduct, it has held the agency responsible.[67] Journalists and courts do the same by finding out what various persons did and said. They may do so by putting their subjects on the record. In both cases, just the creation of a record of what was done by whom secures robust mutual understanding. The record becomes the common account, among citizens, for example, of what happened, and puts a firm foundation under any further thinking about the incident in question, or simply puts the matter to rest conclusively.[68]

---

[66] I think it's probably rare, but possible, to be warranted in asking "You didn't lose those 10 pounds. What happened?"

[67] A good example is the 9/11 report, though, to be clear, that report did find serious mistakes by various government agencies. *See* NATIONAL COMMISSION ON TERRORIST ATTACKS UPON THE UNITED STATES, 9/11 COMMISSION REPORT (2004) (available at https://9-11commission.gov/report/).

[68] In our current political climate, such a basic form of mutual understanding is being eroded, where partisan views of event do not find reconciliation in a common record.

I've been discussing holding persons responsible for non-speech, but persons can be held responsible for their speech too. Considering these cases reinforces my main points, which is that holding responsible puts in place a mutual understanding. This is so because speech is a way of signaling a willingness to be held responsible, by making that task easy. All the other person has to do is listen in order to find out what the speaker wishes to be responsible for.

For example, an audience holds a speaker responsible by listening to what she says, since, in the main, one can believe someone is responsible for what they say.[69] A scholar holds a colleague responsible for what they claim by quoting them, and a journalist holds a politician responsible by listening to, or in some cases recording what they say. Thus, speaking something in one's own name to attentive listeners secures a minimal kind of mutual understanding; namely, the understanding of where the speaker stands on the issue in regard to the community that is made up of those present.

Of course, things don't usually stop there, because in saying something to an audience, the stage is set for further dialogue, in the form of audience's response. He may believe what the speaker says, or confirm that she meant to say what she said, or ask for clarification about the content, or ask her reasons for believing the content that she sincerely asserted. All of these responses are ways of continuing a discussion and of thinking together about what was said, in the hopes of reaching a mutual understanding about it.

Notice that when holding responsible takes the form of hearing what someone asserts as true and asking for its justification, there is no tool, punishment or otherwise, that becomes appropriate to resolve disagreement, as there is about disagreements between persons about whether conduct was justified or not (one person may be legally speaking, in the wrong and so liable to

[69] There are exceptions, such as when one is called to ask "Did you mean to say that?" However, such cases are disruptions to the normal order of things in which one can, by hearing another's words, hold her responsible for saying them and thereby hold her responsible for their content.

various sanctions). When two persons disagree about what is true, neither one may take steps other than continued discussion, plausibly, because there is no basis on which one person can simply insist, as one thinker to another, that her thinking is correct. Instead, the thinking of the two diverge at such a fundamental level that there is no stable ground from which one could be entitled to impose her views on the other. The form that their mutual understanding must take in such a situation is that of an agreement to disagree, which is a placeholder for some hoped-for future reconciliation.

A parting summary may help reinforce the connection between the broad claims I began with and the examples I gave for them. Holding someone responsible for some content is a stretch of dialogical thinking that always begins with a judgment or a belief that they are responsible for it, which is to say, that they can look after it, by speaking for it, answering questions about it, and abandoning it if it's shown to be false. As a judgment, it will be mistaken if the person is not actually responsible for the content.

This judgment is momentous because it is a recognition that one's position must be taken account of in collective thinking. For collective thinking to go forward, the participants must each know where the others stand, and holding responsible is the uptake of where someone stands. But, crucially, holding someone responsible does not just secure a needed requisite for thinking together, but rather, is, itself, the accomplishment of (some) collective thinking, because it secures a minimal understanding, not about what is true or good or creative, but about what the participants think about such things. This mutual understanding may then be relied on or invoked to try for an even greater understanding, because it will warrant the parties in calling on various participants to perform their status as spokespersons for content by calling on them to speak for it, e.g., "What is there to say on behalf of your idea?"

All of these points can be seen in a conversation, which is a microcosmic moment of thinking together in the form of asking to be, and being, held responsible. One person says something, and thereby signals that they are willing to be a spokesperson for it, and the other party hears them, and holds them responsible for what they say by believing they are a spokesperson for it. Already, a mutual understanding has occurred, though it is minimal because it concerns only the content the speaker is responsible for and not the truth or adequacy of it. However, it may be pushed further, because the hearer, on the basis of her correct judgment that the other party can speak for what she said, may ask her to do so, by saying "Why do you think that?" From here, the parties may continue to talk and either reach a more substantive mutual understanding, or else their understanding will take the form of an agreement to disagree, which is a self-consciously tolerant suspension of discussion until a more productive time.

C. Mental Privacy: Shielding Persons from Being Prematurely Held Responsible for Their Thinking

At this point there are two ideas on the table – thinking and holding responsible – and joining them together yields a justification for mental privacy.

Imagine that someone has a thought. She is responsible for its content. After all, it's the product of her mind, and she may revise it when it seems vulnerable to criticism, or maintain it when it seems to survive challenge. However, she's on her own, in the sense that she's on both sides of the dialectic. The decisive criticisms are hers and so are the rebutted challenges (even if she lifts them directly from books she reads). Hence, this kind of solo practice or rehearsal, of trying to see what she really thinks by trying to both buildup and tear down her ideas simultaneously to see what endures. This is thinking. It is an internally constituted exploration or testing out in an attempt to see what this thought is like – what it would be like to champion it, or speak for it, among others. As I argued before, this is an ongoing process.

If another person learns the content of her thought, he will, knowing that persons are generally responsible for their thoughts, have a sufficient reason to believe she is responsible for the thought and thereby hold her responsible for it. This judgment interrupts or preempts the thinking process, because it transforms it from practice into the real thing, the way an audience walking in to a rehearsal makes it suddenly a performance. What was one's working things out, so that one could confidently take responsibility for an idea, is now one's debut idea amid at least one member of the rational community.

Interrupting or preempting a person's individual thinking by holding her responsible for a stage in it, has two objectionable consequences. The first is that it embodies a flawed way of thinking *about her*. It misrepresents her by "freezing" things at an arbitrary point in a dynamic process, where it is the endpoint or resting place that the person chooses for that process that defines her. The second is that the interruption *does something* to her. Specifically, it plunges her into joint thinking when she is not yet clear on what she thinks. The result will be to place various kinds of pressure on her thinking, assuming she becomes aware of the other party's surveillance. I will illustrate these consequences in depth and in order, in what follows.

1. Mindreading as Wrongful Misrepresentation

The starting point for my argument that discovering what someone thinks is a wrongful misrepresentation is the interest persons have in being known, or recognized, as the particular person they are. I think this intuitive. Being misunderstood is a damaging kind of alienation, whereas the value of intimacy partially lies in being fathomed, known, or having one's particularity registered by others, such as by an "understanding" friend.[70] Moreover, racial stereotypes are wrong, *inter alia*,

[70] Lisa Register & Tracy B. Henley, *The Phenomenology of Intimacy*, 9 J. SOC. & PERS. RELATIONSHIPS 467 (1992) (discussing intimacy's relationship to knowing someone). We sometimes say "I'm thinking of you." This locution is interesting because it conveys solidarity, and it appears to refer to a way of thinking about a person, but not under any

because they force persons of a certain race into an inaccurate and homogenizing social identity.[71] Relatedly, it's demeaning to refer to someone as a number (such as in a prison), or with a unilaterally chosen nickname, as these are attempts drown out or dodge a person's particularity.[72] Correspondingly, a familiar, perhaps vengeful, demand for recognition is "Say my name!"

This interest in being recognized is of course violated by refusing to acknowledge a truth about a person,[73] or by spreading lies about them, but it is also violated by learning truths about a person that compromise one's ability to go on thinking fairly about her. Information like this, that is true but disruptive to one's fair-mindedness, is prejudicial (etymologically, pre-judging). I will argue that a person's thoughts are a quintessential source of prejudgments about her.

A brief elaboration of the nature of prejudice is helpful. For my purposes, prejudicial information is a truth about something that diminishes one's ability to reason further about it, in the way that tasting a strong flavor may diminish one's ability to appreciate other, more subtle flavors in the same dish. Put differently, prejudicial information is epistemically volatile. It is a truth, but one that has "explosive" – unpredictable and uncontrollable – consequences for the thinking of one who learns it, so that it becomes difficult for them make further judgments in a way that is commensurate with the relevant reasons. Prejudicial information is thus aptly named because it encourages persons to jump ahead to conclusions about someone before completing the intervening activity of really getting to know them.

---

label. That is, it seems to be a way of thinking about someone simply as the particular perspective they are, unburdened by their contingent characteristics, such as their age or sex.

[71] JAMES BALDWIN, THE FIRE NEXT TIME (1963) (arguing at a number of points that a deep problem of racism is the misperception (and outright fabrication) of the identity of the subjugated race.) So-called positive stereotypes are no less damaging because they are just as much a misrepresenting pre-judgment of those who are subject to them.

[72] Mariana Souto-Manning, *Challenging the Text and Context of (Re)Naming Immigrant Children: Children's Literature as Tools for Change, in* PROMOTING SOCIAL JUSTICE FOR YOUNG CHILDREN 111 (Beatrice S. Fennimore & A. Lin Goodwin eds., 2011) (discussing the disregard for identity involved in renaming school children).

[73] For example, parents who refuse to acknowledge that their child is gay, or newspaper articles writing about clearly gay men as "lifelong bachelors." *See* LARRY GROSS, CONTESTED CLOSETS: THE POLITICS AND ETHICS OF OUTING (1993).

A stark example is racial information. In a society where racial stereotypes are present, knowing a person's race may provoke judgments about her that are not at all supported by the fact that she is a certain race.[74] Such judgments are nothing more than racist ideology. However, some information is more ambiguous in that it may provide some genuine support for a conclusion, but does so in such an inchoate or hard to grasp way that it cannot be properly incorporated into one's thinking. In such cases, balancing is needed to determine whether it is predominantly probative or prejudicial. Courts are regularly called to perform this kind of balancing, and have developed rules of thumb and canons of reasoning to assist them in doing so.[75] For example, information about a defendant's past criminal convictions – her *record* – is usually excluded from a trial for a different crime, on the grounds that such information, though somewhat probative of whether she committed it, is too likely to "over-persuade" the jury that she is guilty by painting her as simply, inchoately, "bad."[76]

A second useful illustration is art. An artist working on a sculpture is engaged in aesthetic thinking, with the sculpture being that thinking, or working, in progress and embodied. Like thinking generally, the artist's work has no determinate ending point.[77] It is brought to an end by a decision that the work is done, and consequently, an overzealous critic who views the work before

---

[74] The solution to this problem may sometimes be to prevent information about a person's race from becoming known. *See, e.g.*, David Hausman, *How Congress Could Reduce Job Discrimination by Promoting Anonymous Hiring*, 64 STAN. L. REV. 1343 (2012).

[75] *See* Fed. R. Evid. 403 ("The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.").

[76] *See* Fed. R. Evid. 404(b) (prohibiting introduction of past crime to show that alleged crime was committed); Michelson v. United States, 335 U.S. 469, 475–76 (1948) (past convictions "overpersuade" the jury); United States v. Moccia, 681 F.2d 61, 63 (1st Cir. 1982) (discussing the rationale for the rule against introducing past crimes to avoid prejudice).

[77] *See, e.g.*, Cour d'appel [CA][regional court of appeal] Paris, Mar. 19, 1947, D.P. 20 (*L'affaire Rouault*) (art dealer kept paintings that an artist would modify from time to time, and it was found that ownership had not passed to the dealer because the paintings had not been finished). *See also* Carl Settlemeyer III, *Between Thought and Possession: Artists' "Moral Rights" and Public Access to Creative Works*, 81 GEO. L.J. 2291 (1993) (quoting *L'affaire Rouault* "the painter remains master of his work, and may perfect it, modify it, or even leave it unfinished if he loses all hope of making it worthy of himself.").

that time injures the artist.[78] By seeing the art before the artist certifies it as finished, the critic gets a distorted glimpse of it and so, also, the artistic personality at work in it.

Moreover, the distortion is prejudicial, because it casts a shadow over subsequent thinking that is hard to counter. Once the critic has seen the piece in progress, it's hard to unsee, and so hard not to see the finished piece *through* the earlier glimpse of its draft. The art, in virtue of being seen unperfected, cannot "stand on its own" as an object of consideration, but will share mental space with its superseded version.[79] I hope the connection to my subject matter is clear: persons are the products of thinking about what is true in the way that a piece of art is the product of aesthetic thinking, and they have a similar interest in standing on their own as an object of thought, which they cannot do if they are dogged by their experimentation.

Thoughts are, like criminal convictions or the state of an unfinished creative endeavor, always or almost always more prejudicial than probative, and for that reason informational mental privacy categorically blocks access to them. Some subtleties are worth remarking on. One is that legal evidence rules and artistic decorum are for the benefit of specific persons tasked with making specific kinds of judgments, but informational mental privacy sweeps more broadly, because its goal is not to secure a fair trial, but something more ambitious: the entrance of persons into social life on terms that are favorable for their being understood as the persons they are.

A second subtlety is that a past conviction risks prejudicing a jury because of the tendency for persons to overreact to *wrong*doing, whereas thoughts risk prejudice because of their rightful

[78] *See* Raymond Sarraute, *Current Theory on the Moral Right of Authors and Artists Under French Law*, 16 AM. J. COMP. L. 465, 467 (1968) (summarizing artistic rights under French law: "So long as a work of art has not been completely created—of which the artist alone can be the judge—it remains a mere expression of its creator's personality, and has no existence beyond that which he tentatively intends to give it."); Jon Newman, *Copyright Law and the Protection of Privacy*, 12 COLUM.-VLA J.L. & ARTS 459 (1988).

[79] *See* E.M. Forster, *Anonymity: An Inquiry*, 2 CALENDAR OF MODERN LETTERS 145, 150 (1925) ("[Literature] is always covering up the tracks that connect it with the laboratory.").

centrality to persons.[80] As I mentioned before, persons have an interest in being known, but they also have a symmetrical interest in knowing others and building intimacy with them, and because this is most fundamentally achieved by knowing what they think, thoughts rightly arouse the attention and concern of persons. We want to know what others think to learn from them, to appreciate them, and to build a shared understanding of how to live that nonetheless respects, as much as possible, each person's divergent viewpoints.[81]

The problem is that thoughts are liable, precisely because of how important they are, to attract too much attention and to be too interesting, so that persons can become too voracious in seeking them out.[82] Celebrity culture and the faux-intimacy it thrives on is a good example, as well as the misuse of social media to indulge in sensation, triviality, and narcissism.[83] This kind of uncontrolled interest is aroused by even rough, unrefined thoughts, because a bit of a person lives even in her flights of fancy and exploratory judgments. All thoughts are, to some extent, indicative of who one is, but the mistake of those to whom mind reading appeals is in supposing that to know someone well involves taking account of even these idiosyncratic and incomplete expressions of personhood. Perhaps they even think that the rawness of such moments should be equated with authenticity.[84]

---

[80] Unfinished art is prejudicial in a way that is more like thoughts in this regard, because its prejudicial impact stems from the rightful interest that persons have in engaging with the creative work of others.

[81] ANTHONY LADEN, REASONING: A SOCIAL PICTURE (2012); David Beglin, *Two Strawsonian Strategies for Accounting for Morally Responsible Agency*, 177 PHIL. STUD. 2341 (2020).

[82] Louis Menand, *The Lives of Others,* NEW YORKER Dec. 25, 1969; DAVID LODGE, CONSCIOUSNESS AND THE NOVEL: CONNECTED ESSAYS (2002) (arguing that the novel as an art form is compelling and distinct partly because of how it gives persons access to the thoughts of the characters).

[83] CLAY CALVERT, VOYEUR NATION: MEDIA, PRIVACY, AND PEERING IN MODERN CULTURE (2000); David Bromwich, *How Publicity Makes People Real*, 68 SOC. RES. 145-171 (2001); Barry King, *Stardom, Celebrity, and Para-Confession*, 18 SOC. SEMIOTICS 115 (2008).

[84] *See* Rachels, *supra* note 6, at 290-1 (recounting a story in which some of the characters seem to believe that how a person acts when no one is looking is who they really are).

But the opposite picture is closer to the truth. As I argued before, thoughts are not aspects of a fully formed person, but the scaffolding out of which a fully formed person will be constructed, and so trying to learn about someone by reading her mind is folly. Thinking follows no predictable path, and what someone thinks today is an unreliable guide for the position she will end up settling on tomorrow. Moreover, being exposed to a person's thinking pushes one toward judgments about her on the basis of something that is only true of her-under-construction. This is a literal kind of prejudgment, of evaluating the product on the basis of one moment in the working-out of its creation, not unlike judging a novel on the basis of one of its earlier drafts.[85]

To give a concrete example, imagine that it is somehow revealed that one's neighbor believes the moon landing was faked, or that vaccines are dangerous, or that persons should be allowed to sell their organs, or any number of other viewpoints. One will likely start to think of the person, involuntarily, as owing allegiance to that content, when in fact he only thinks it. Moreover, one may also have a hard time resisting the collateral conclusion that he is, on the basis of such an allegiance, gullible, unscientific, mercenary, and so on.

It may seem that the problem I'm pointing to is overblown. Won't many thoughts be banal and so provoke no prejudice? And can't one remind oneself that a thought is just a thought, so that when the neighbor comes to talk about politics, one will not hold his moon landing views against him? Perhaps, but my claim is not that reading someone's mind is wrong because it always produces a rabidly prejudiced state of mind. Rather, my claim is that reading someone's mind is wrong because it always unjustifiably risks prejudice, just as it's wrong to drive without lights at night

[85] JEFFREY ROSEN, THE UNWANTED GAZE 8, 9, 12, passim (2000) (discussing how privacy violations misrepresent persons, but because they reveal information out of context – or as just a part of a whole).

because that nearly always unjustifiably risks an accident, even if, in a specific case, no accident occurs.

In the case of mind-reading, the risk of prejudice is unjustifiable. Many thoughts that one might discover will provoke strong reactions. One can try to resist them in the name of fairmindedness, but doing so requires vigilance, which may flag, or be bypassed by unconscious or emotional alterations to one's objectivity. Moreover, resistance itself is fraught, because it can easily go too far (if there is even a matter of fact about how far it should go). In trying to overlook someone's belief in a baseless theory, one may go too easy on their other mistakes and flaws, for fear it is one's prejudice talking.

It's true that the foregoing pressures may be absent when one uncovers only ordinary thoughts by mindreading, but the risk of prejudice is still there. After all, it's hardly possible to limit one's mindreading to only banal thoughts, and what is banal to one may be inflammatory to another. Moreover, even finding out what banal thoughts someone has may be prejudicial. It may give one the impression that the thinker is dull or conventional, which may lead one to interact with them differently, by avoiding them as uninteresting, or to defer excessively to what they say about other issues, on the assumption that their viewpoint is similar to one's own. The point is that even someone with conventional beliefs has an interest in being able to be known for what they do and say, and not what they happen to believe. Persons have an interest in being able to control the unfolding of their own life narrative, so that, as far as possible, they do not have to compete with their past thinking in being accurately known. This is an interest in being able to stand on their own as objects of thought.

2. Mindreading as the Wrongful Initiation of Thinking Together

According to the argument I just gave, prematurely holding a thinker responsible for her thoughts risks a bad psychological consequence, that of prejudicing one's view of her. Crucially, the argument does not assume that this psychological consequence invariably follows from premature mind-reading. Rather, it claims that premature mind-reading constitutively risks this psychological consequence, in the sense that learning what someone thinks is, constitutively, the start of thinking about her, even if, in a particular case, the mindreader manages, through luck or force of will, to keep himself from jumping to any prejudiced conclusions. This is similar to the way in which insulting someone constitutively risks making them feel bad, even if, in a particular case, they shrug off the insult.

The argument of this section has a similar structure. I will argue that not only does mind-reading constitutively risk thinking *about* the thinker in the wrong way, but it also constitutively risks *doing something* to her. This risk stems from the fact that knowing someone's thoughts is not only the start of thinking about her, but the initiation of thinking with her, in a roughly dialogical way. And dialogue constitutively forms a compound or reciprocally mediated perspective on a subject matter. The effects of prematurely being made a party to such a perspective come in the form of distortive, purely psychological parodies of the agreements and disagreements that accompany ordinary conversation: conformity or non-conformity. I explain this argument by first looking at the structure of dialogical thinking embodied in a conversation, and then linking it with psychological research, most of which focuses on the way overbearing surveillance produces conformity (rather than non-conformity) in subjects.

Conversations are joint-thinking on a small scale, and yet they set in motion powerful rational effects, chief among them being the creation of a joint perspective on the subject matter

being discussed. This standpoint is constituted by the intermixing of the individual thinking of each conversant, in the way that a major river is constituted by two tributaries that join to meet it, and in the way that two contractual parties make themselves parties to a single perspective of what the course of their dealings together will look like.[86] Conversants create and make themselves parties to a joint perspective on the topic of conversation, consisting of a fluid, shared set of assumptions that linguists often study under the heading of a "common ground."[87]

When we believe what we are told, the conversational joint perspective is lopsided. It consists of one party signing on wholesale to the other's thinking. At the other extreme, when the parties agree to disagree, the thinking of both parties remains distinct and unfused, though mutually understood to be that way. In between these two extremes, the conversants start from different positions and meet somewhere in the middle, with the result being an amalgamated joint perspective that reflects both of their contributions in the course of the conversation.

The creation of a joint perspective is what gives conversations their great power, and often enough, their great value, but it also poses dangers. The reason is that both the process of creating a joint perspective and the resulting perspective itself place significant demands on a person's thinking.

The creation and maintenance of a joint perspective consists in encountering, and seeking to accommodate the view of the other, to whatever degree is reasonable. Doing this breaks down into a series of rational maneuvers, involving the articulation of one's viewpoint, understanding that of the other, insisting on a way of seeing things, pressing a metaphor or point of focus, but also yielding to the better argument, going with the flow of conversation, and entertaining other ways of

[86] This way of talking can be found in reasoning about contracts. *See, e.g.*, AT&T Commc'ns, Inc. v. Wiltel, Inc., 1 F.3d 1201, 1206 n.3 (Fed. Cir. 1993) (reasoning at one point "from the perspective of the contract as a whole").

[87] Robert Stalnaker, *Common Ground*, 25 LINGUISTICS & PHIL. 701 (2002) (the locus classicus).

viewing the subject matter. I am not suggesting that all conversations are power struggles or conflicts. They are not. Instead, my point is that they are very often cooperative, and that when they are, the cooperation that they call for requires skill and effort. They require mobilizing one's view of what is true, interesting, illuminating, worth saying, while at the same time being open to a different view of those same things. Broadly speaking, the dynamics of a conversation are pushes and pulls, and navigating them is a matter of sound judgment and receptivity to the thinking of another.

The joint perspective the emerges from these rational maneuvers likewise has rational consequences. It becomes a part of the thinking of the conversants. It frames their understanding of a subject matter and suggests new angles on it. Much thinking involves imagining what others would say or recalling what they have said, and conversations routinely impress us with their ability to present things from another perspective, and thereby knock us out of our own comfortable cognitive orbits.

Crucially, whether the demands that conversation places on thinking are constructive depends on the readiness of the parties to meet them, and a crucial part of being ready to meet them is having done some thinking of one's own first. The difference this makes is in how it structures conversational demands. The pushes and pulls of conversations are supposed to be exercises of one's own thinking, and the resulting joint perspective is supposed to advance it. Consequently, if one has done some thinking already, the predicate of these demands is in place, and they can be satisfied as the exercise of one's thinking and its advancement.

By contrast, if one has not thought about an issue to one's satisfaction, then the push and pull of a conversation about it and the resulting joint perspective will not exercise or advance it. Instead, the pushes and pulls of conversation will take the form of psychological pressures; rather than agreeing or disagreeing, one will be conforming or non-conforming. And the resulting joint

perspective will not meet with and transform one's own perspective, but rather will supplant it, mediate its contact with the relevant subject matter, and distort it. In short, participating in a productive conversation requires sufficiently settling one's own thinking first, or else sufficiently trusting the other to respect one's unsettled state (which may be easier when both parties are undecided and wish to think through an issue together). Without these requisites, only a faux-conversation is possible, and the processes and product that ordinarily constitute it (a joint-perspective) will be faux versions too.

A compact summation of these points is that a conversation goes well when both parties can encounter the ideas of the other through, or as a response to, or an elaboration of, their own direct, solitary encounter with the relevant reasons. If one has not engaged in this unmediated thinking first, then the views of others are not proposals or suggestions, but de facto starting points. By contrast, by staking out a position, one has an orientation from which to confront other arguments and perspectives. The old adage, attributed to Alexander Hamilton, puts it artfully, "If you don't stand for something, you will fall for anything."[88]

A number of examples support this idea. In one version of the Asch experiment paradigm, subjects were asked to answer a question after a number of confederates answered it incorrectly. The false answers of the confederates induced some degree of conformity on the part of subjects leading them to give false answers too. However, such conformity was notably reduced when subjects were given a chance to write out their answer first, and thereby "pre-commit" to it.[89] Other psychological studies indicate that when persons expect to have a conversation with someone, they will often

[88] Standing for something is a part of realizing the virtue of integrity. *See* Calhoun, *supra* note 20.

[89] Morton Deutsch & Harold B. Gerard, *A Study of Normative and Informational Social Influences Upon Individual Judgment*, 51 J. ABNORMAL & SOC. PSYCHOL. 629, 634 (1955) ("…the more uncertain the individual is about the correctness of his judgment, the more likely he is to be susceptible to social influences in making his judgment.").

conform their viewpoint toward theirs.[90] In some cases though, such as when persons perceive the revelation of their thoughts to be unfair, they may cling to them more stubbornly.[91] In these cases, the lack of preparation for a dialogue transforms what might be agreement or disagreement into pressure for conformity or non-conformity.

Ordinary examples are supportive as well. In-person conversations initiated by one party may be excessively influential, which is the rationale for rules against prohibiting lawyers from soliciting clients in person.[92] George Tsai gives several examples that illustrate how persons who have not made up their own mind may be excessively vulnerable to the viewpoints of others. In one example, he notes that scholars undertaking a research project may purposefully shield themselves from the viewpoints of other scholars in the field so as to get a grip on what they think first, before seeing how it matches up with the approaches of others.[93] In a different example, he considers a political philosophy professor who shields his true views from his students so not to bias their own thinking about the issue.[94] One could imagine a version of the example in which the professor did tell a graduate student his view on some issue. If the graduate student had not thought about the issue much first, the comment could have an outsized affect on her deliberations and research path. She may also wonder, if her own interest and arguments on the subject are really her own, or really belong to the professor.

---

[90] Jennifer Lerner & Philip Tetlock, *Accounting for the Effects of Accountability*, 125 PSYCHOL. BULL. 255, 256-7 (1999). *See also*, Margot Kaminski & Shane Witnov, *The Conforming Effect: First Amendment Implications of Surveillance Beyond Chilling Speech*, 49 U. RICH. L. REV. 465 (2015) (discussing the psychological research about conformity).

[91] *Id.* at 258-9.

[92] MODEL RULES OF PROFESSIONAL CONDUCT R. 7.3(b) (limiting in-person solicitation of clients); Ohralik v. Ohio State Bar Ass'n, 436 U.S. 447, 457 (1978) ("in-person solicitation may exert pressure and often demands an immediate response, without providing an opportunity for comparison or reflection"); McDaniel v. Pettigrew, 536 S.W.2d 611, 614 (Tex. Civ. App. 1976) (reviewing law that provided for a cooling-off period for contracts entered into as a result of in-person solicitation at a person's residence).

[93] Tsai, *supra* note 32, at 99.

[94] *Id.*

To close this section, I will tie the argument back to mind reading. Mindreading holds a thinker responsible for their thoughts, and holding someone responsible is, in principle, the start of having a dialogue with her. This is morally objectionable because dialogues aim at creating a joint perspective, and being a party to a joint perspective is damaging to one's thinking if one is not prepared to play that role. The damage may take the form of serious psychological pressure toward conformity or non-conformity.

This is not to say that each and every instance of mindreading will impose significant psychological pressure on the subject. We might not be aware that we are the victim of mindreading, or we might be the victim of mindreading by an unknown person, or by someone we will never see or converse with, or by someone whose opinion does not matter to us. In these cases, we are unlikely to alter or abandon our course of thinking. The damage will be negligible or non-existent. However, such cases of mindreading are still wrongful, and the reason is that they initiate a dialogue, when doing so constitutively risks harm to the target's thinking, even if, as things are, it is unlikely to actually cause it.

3. The Relationship Between the Two Justifications

To close this section, I want to briefly discuss the relationship between these two justifications, by asking, which one is fundamental? Many authors write as if violating mental privacy is wrong because of the effect it has on the target's thinking itself, rather than because of the effect it has on the violator's thinking *about* the target.[95] To their credit, this prevalent view notices that being aware that one is the victim of mindreading unbalances one's thinking, but it struggles to say why covert mind reading is a wrongful violation of mental privacy, because the thinking of the victim need not be affected by someone who pulls it off without alerting her. A person whose thoughts

[95] Kaminski & Witnov, *supra* note 90; NEIL RICHARDS, INTELLECTUAL PRIVACY 104, 106 (2015).

have been read but does not know it will keep thinking as before. A second, related problem is that some psychological studies suggests that the damaging effects that overt mindreading has on thinkers depends on their belief that a particular instance of it is wrongful, which means that some of its psychological consequences on the victim are a symptom and not the ultimate basis of its wrongness.[96]

It may seem that my view should be the opposite; namely, that mindreading is primarily wrong because it misrepresents the thinker, and that, in the case where the mindreading becomes known, further collateral damage may be done to the target in the form of psychological pressure. Such a view is plausible. It straightforwardly illustrates what is wrong with covert mindreading; namely, that is misrepresents the thinker, which is a harm that she suffers, but not through any awareness on her part. Moreover, it would explain how some kinds of distortion of thinking that arise from mindreading could be the consequence of the target's awareness of being wrongfully misrepresented.

Still, I think an even better position is that neither misrepresentation nor distortion is the fundamental wrong of mind-reading, but rather, that they are symptoms of a more general wrong, which is the wrong of officiousness, or, of making someone else's "business" partly one's own by "minding" it. For example, one of my friends should not plan and execute my daughter's birthday party. Doing that is something that I, along with my partner, am responsible for, but not only that, it is something that I have an interest in being responsible for with her, as it's a task that demonstrates and constitutes our special role in our daughter's life. It is, in the first instance, our business (though of course we can ask for help).

[96] Lerner & Tetlock, *supra* note 90, at 258-259.

The same is true of thinking. Persons are responsible for their thoughts, and they have an interest in being solely responsible for them, insofar as they are still working them out. This is a matter of getting a chance to be in touch with ideas, whether truths or fictions, in a way that is unmediated by others. Holding someone responsible for something is, as I have argued, the other side of things. It is an attempt to structure communal thinking by first taking account of what someone thinks and so to make their business into the community's business. At certain point, concern for how others conduct themselves becomes morally urgent, and good. Before that point is reached though, lies meddling. Hence, learning a person's thoughts is a consequential kind of usurpation or take-over, even if only in intent.[97] It is a failure to mind one's own business.

III. Conclusion: Resolving the Dilemma and Other Kinds of Privacy

To conclude, I return to two claims I made in the introduction. First, I hypothesized there that the purpose of privacy is to prevent a person from being held responsible. This idea seems to founder on a dilemma: either privacy allows persons to avoid being held responsible for their failures to live up to sound norms, or else it is an odd kind of stopgap measure whereby the full effects of intolerance or immorality are blunted, but their roots are left in place pending a more thorough reform of attitudes and institutions.

I argued for a middle way, according to which holding persons responsible has dangers if practiced excessively, and where the excessiveness is not a matter of being pursued inefficiently or wastefully, such that there are more pressing matters for persons to pay attention to than the thoughts of others. Rather, the excessiveness is a matter of holding persons responsible for aspects

[97] *Cf.* Seana Valentine Shiffrin, *Paternalism, Unconscionability Doctrine, and Accommodation*, 29 PHIL. & PUB. AFF. 205 (2000).

of themselves that are so nascent and vulnerable that they are not ready to withstand the former. Holding persons responsible for thoughts is, *per se*, an overuse of holding persons responsible.

If we accepted the foregoing, but thought that holding responsible was only a matter of imposing consequences on persons, then we would be forced to conclude that mental privacy is not violated merely by finding out what someone thinks, because doing so may not impose any such consequences. Instead, I think we should see the significance of holding responsible as owing, fundamentally, to the significance of being exposed to the attention of others. This exposure itself is significant enough, indeed more so than consequences, because a person is not the kind of thing that is what it is, independent of the way it lives in the thinking of others. Instead, a person emerges from a history of interpersonal moments and maneuvers, in the way that a corporation is (an artificial person) built out of contracts, handshakes, filings, declarations, and the assumption of rights and responsibilities. For persons, publicity (whether between just two or many) is reality, because persons are constituted by thinking, and thinking is essentially, really, or truly, a public activity, carried out between persons aware of each other's thoughts, though it may be practiced in isolation from others, by installing oneself on both sides of an interpersonal dyad. In thinking by oneself, one is setting up the positions one will and will not make public, which will in turn determine who one asks to be, and if, recognized, who one really is, as a community member. To be rushed into taking form is a gross violation and the one that mental privacy guards against.

This view is not the editorial freedom justification all over again. The justification for giving persons a chance to form themselves is not the interest persons have in being the way they want to be. Instead, the view I have put forward is bounded by the needs of cooperative and social life. That is, it acknowledges the need for persons to give themselves a core of self-thinking before being asked to do what is, essentially, to consider others, and often enough, to transform themselves to

accommodate them. The entitlement to think, by oneself, is fundamentally secured by the importance of being correctly sensitive to group concerns, in the form of group deliberation.

This connection to group deliberation builds in a responsiveness to others at the fundamental level, whereas the editorial view highlights an interest in privacy that extends only to the border of each person's individuality, and provides no framework for those interests to interact. Some fences are just for keeping others out, and the editorial view treats privacy this way, as a right to exclude others so as to serve one's self-presentational interests. However, some fences, as the saying goes, make good neighbors, and on the view presented, mental privacy is such a fence, because the separation it creates paves the way for a more respectful and productive kind of togetherness.

The second loose end from the introduction is my claim that mental privacy has an important place in understanding privacy more generally. If this result could be shown, it would unsettle a theme of contemporary privacy scholarship, which is that privacy is a grab-bag concept with no unifying purpose or justificatory structure.[98]

Some hope for this idea lies in the application of what I have argued to other kinds of privacy. After all, a self is more than thoughts. It consists of categories that one identifies with (race, religion, sexual orientation, gender, etc.), a life-history of what one has done, and the body that one inhabits. Though these things are not thoughts, they are the object of a person's thinking in distinct ways, and correspondingly the object of distinct thinking by others. For this reason, persons often have a need to think about them first, to reach a self-understanding before putting them forward as potential points of mutual understanding.

[98] DANIEL SOLOVE, UNDERSTANDING PRIVACY (2008) (arguing that privacy labels a hodgepodge of different moral interests).

Privacy lives where this order needs to obtain; when self-understanding needs to precede collective understanding. A simple example is sexuality. Persons often need to think about what sexual orientation they wish to "come out" as having, and this requires experimentation. Some experimentation may be mental, in the form of processing feelings of attraction and love, but some experimentation may need to take the form of intimate actions. In this case then, we see different kinds of privacy protecting different stages in a person's journey to self-understanding. Mental privacy may be needed to think about one's sexual orientation, privacy of the home and of intimate spaces may be needed to have intimate physical contact with whom one chooses, and identarian privacy may be needed to protect the tentative conclusions one reaches about who one is. Much more needs to be said, but it seems that there is a richness to the ways in which persons get ready to be thought about by others, and hence, a richness to the forms that privacy may take.

Chapter 2: Anonymity

I. Introduction

In the preceding chapter, I argued that persons signal their readiness or unreadiness to think with others by revealing or withholding their thoughts. Because of the risks of initiating thinking with someone who is unready to do so, such signals should generally be respected. However, there is another way of signaling unreadiness to think with others besides withholding one's thoughts. Instead, one can volunteer the thoughts, but withhold, as it were, oneself as the thinker of those thoughts. This is accomplished through anonymity.

Anonymity can mean the unrecognizability of a body, such as when a person wears a mask, but it can also refer to the namelessness of a speaker or a reader. Both are important, but I focus on the latter: content that is not attached to a name. This kind of anonymity makes an appearance in all sorts of places, and it alters the dynamics of any activity in which it crops up. Publishing a book in one's own name is different than publishing it anonymously;[1] reporting a crime in one's own name is different than giving an anonymous tip;[2] and signing an opinion in one's own name is different than joining a *per curiam* opinion.[3] Not surprisingly then, given its ubiquity and significance, anonymity is the subject of a number of legal doctrines and practices. These practices are controversial, as

[1] Many authors publish pseudonymously. Pseudonyms introduce complications that I will deliberately leave aside in this chapter, but some uses of it are illustrative, since someone who uses a pseudonym does not speak in their own name. *See* Joseph Fulda, *The Ethics of Pseudonymous Publication*, 16 J. OF INFO. ETHICS 75 (2007). *See also*, Alexandra Schwartz, *The "Unmasking" of Elena Ferrante,* THE NEW YORKER Oct 3, 2016 (discussing the identification of a popular novelist who had written under a pseudonym).

[2] Navarette v. California, 572 U.S. 393 (2014) (holding that an anonymous tip may satisfy the requirement of reasonable suspicion when sufficient indicia of reliability are present, such as the identity of the car that was reported to have driven dangerously).

[3] *E.g.,* Ira P. Robbins, *Hiding Behind the Cloak of Invisibility: The Supreme Court and Per Curiam Opinions*, 86 TUL. L. REV. 1197, 1199 (2012) (arguing that courts have reason not to speak as institutions and should, generally, speak through the opinions of particular judges).

evidenced by the deluge of articles that take up the case for and against anonymity in areas ranging from civil and criminal procedure,[4] family law,[5] constitutional law,[6] employment law,[7] and more.

This body of work has charted a vast territory of anonymity doctrines and a diversity of considerations that bear on whether to retain or reform them. However, the impressive breadth of this scholarship leaves some of the depths unplumbed, because it is built on an uninterrogated assumption about the kind of thing anonymity is. This assumption is that anonymity is just a condition in which others are ignorant about one's identity, where such a condition is no different, in principle, than parallel conditions regarding other facts, such as about one's blood type or true hair color.

To illustrate the nature of this assumption, compare a mundane tool, like a hammer, with a moral "tool," such as punishment. Both are similar in that they have a purpose or something that they are for. A hammer is for driving nails and punishment is for holding persons responsible for wrongdoing. Notice though that their respective purposes differ in their moral significance. Using a hammer for the purpose of driving nails is neither necessary nor sufficient for using it in a morally justified way. It's not morally wrong to use a hammer as a paper weight, and it is morally wrong to drive nails into someone else's house without consent. The point is that whether a particular use of a hammer accords with its purpose of driving nails makes no moral difference, in itself.

---

[4] *E.g.,* United States v. Mansoori, 304 F.3d 635, 649 (7th Cir. 2002) (analyzing the propriety of empaneling an anonymous jury); United States v. El-Mezain, 664 F.3d 467 (5th Cir. 2011) (analyzing the propriety of anonymous witnesses); Sealed Plaintiff v. Sealed Defendant, 537 F.3d 185, 188 (2d Cir. 2008) (case of first impression discussing the propriety of allowing a pseudonymous complaint).

[5] *E.g.,* Inmaculada De Melo-Martin, *The Ethics of Anonymous Gamete Donation: Is There a Right to Know One's Genetic Origins?*, HASTINGS CENTER REPORT, Mar.-Apr. 2014, 28.

[6] James Gardner, *Anonymity and Democratic Citizenship*, 19 WM. & MARY BILL RTS. J. 927 (2011); Lyrissa Barnett Lidsky & Thomas F. Cotter, *Authorship, Audiences, and Anonymous Speech*, 82 NOTRE DAME L. REV. 1537, 1538 (2007).

[7] David Hausman, *How Congress Could Reduce Job Discrimination by Promoting Anonymous Hiring*, 64 STAN. L. REV. 1343 (2012).

Punishment is otherwise. Using it for its moral purpose *is* necessary and sufficient for its use to be justified. It's morally objectionable to punish a person if they did not commit a legally recognized wrong, even if doing so would avert a riot.[8] That is an abuse of punishment. And it is permissible to punish a wrongdoer, so long as various other aspects of a fair justice system are in place.[9] That is a proper use of punishment. The fact that punishment is susceptible, where a hammer is not, to being *properly* used or abused in a morally significant sense, owes to punishment having a designated moral purpose. We can elaborate this point by saying that the nature of punishment – holding persons responsible for legally recognized wrongdoing – provides for, or determines, the conditions under which its use is justified. Punishment has its own, distinct, moral logic.[10] It can't be properly used to pursue any goals one wants, not even the efficient production of a bundle of important goods, because it is reserved for addressing a distinct kind of moral need.

The assumption of many commentators, often implicit but voiced as well, is that anonymity is as morally mundane as a hammer; that it is a tool in the ordinary sense,[11] without any moral logic of its own, and hence may be used so long as doing so will be sufficiently productive of various morally or legally recognized goods. Specifically, the going view assumes that just as a hammer is for driving nails, anonymity is for producing or maintaining ignorance about the identity of a speaker,[12] and just as driven nails can be useful for all sorts of things, so ignorance can be useful for

---

[8] This example is a standard counterexample against utilitarianism. *See* H.J. McCloskey, *An Examination of Restricted Utilitarianism*, 66 PHIL. REV. 466 (1957) (one of the first presentations of this example).

[9] The moral logic to this side of punishment, the sufficient conditions, is complicated and there is room for many approaches. *See, e.g.*, David Dolinko, *Some Thoughts About Retributivism*, 101 ETHICS 537 (1991).

[10] For instance, punishment must be for a wrong, announced in advanced, proportional to the wrong, and evenhandedly applied.

[11] *E.g.,* A. Michael Froomkin, *Lessons Learned Too Well: Anonymity in A Time of Surveillance*, 59 ARIZ. L. REV. 95, 145 (2017) ("Communicative anonymity is a core part of freedom in a democratic state and a critical tool for those who seek freedom from nondemocratic states."); Bryan H. Choi, *The Anonymous Internet*, 72 MD. L. REV. 501, 538 (2013) ("There is good reason why anonymity and generativity are key pressure points. Both are tools…"); Saul Levmore, *The Anonymity Tool*, 144 U. PA. L. REV. 2191 (1996).

[12] Or reader, or thinker, and so on. I focus here on the speaker, because it is foremost in the mind of many commentators. I discuss readers in section II.B below.

incentivizing various goods, such precluding wrongful action by others, and so on. This assumption paves the intellectual path for a case-by-case assessment of the various goods that anonymity implicates. I call this the assumption that anonymity is *not (morally) distinct*.

I contend that this assumption is false. Anonymity is more like punishment, and so, morally distinct. The parallel can be summarized like this. Punishment is morally distinct in that it does, by its nature, something that calls for justification: it imposes burdens on persons. Hence, the proper use of punishment is one that, at least, warrants the imposition of those burdens. Likewise, anonymity does, by its nature, something that calls for justification: it allows persons to avoid being held responsible for the ideas that they associate themselves with. Hence, the proper use of anonymity is one that warrants persons in avoiding responsibility in this way. I will claim that this use is that of participating in communal thinking, when one's own thinking is not firm or still in progress, such that one is unready to be held responsible for it. Or put another way, if the proper use of punishment (and so imposing burdens) is as a response to wrongdoing, then the proper use of anonymity (and so avoiding responsibility) is as a *privilege to participate in community thinking*.[13]

After arguing for this theoretical claim, the *privileged participation view* for short, I show how it reshapes the kind of arguments that can and should be given for various democratically important uses of anonymity. First, I consider anonymous speaking and reading. They are comparatively straightforward applications of the privileged participation view. After all, persons may need, to advance their own thinking, to make contact with ideas and arguments that are found in a thinking community without yet taking public responsibility for a position within that community. Hence, the

[13] *See, e.g.*, JANNA MALAMUD SMITH, PRIVATE MATTERS: IN DEFENSE OF THE PERSONAL LIFE (1997) ("Anonymity, you might say, is privacy for people who don't want to be really alone.").

use of anonymity to champion dissident positions or read disfavored literature are proper uses. They realize the purpose of the privilege that anonymity embodies.

Other democratic practices are more complicated. I consider three democratic practices: citizen voting (in elections), voting by legislative representatives, and signing a ballot initiative, by which a particular law is submitted to directly to the citizens of a polity for acceptance or rejection. Democracy is a particular kind of community of thinking. One important aspect of it consists of a pattern in which there is a consideration of a question by citizens, whose answer to it constitutes the starting point for further consideration of the question by other public officials who then take responsibility for their answer to it. For instance, the people, in the form of a grand jury may be asked whether an illegal act occurred, and how they answer this question will help determine what public acts officials such as prosecutors and judges take responsibility for, in response (e.g., bringing charges, convening another grand jury).

Elections are similar. Citizens voting in an election are asked to consider which of a series of choices serves the public good. They have an interest in giving an answer, as this is a fundamental means of participating in government, but they may not give an answer if doing so means taking a public position on a question that they are not done thinking about. Secret balloting removes this source of hesitancy. Citizens are freed to vote their best guess about what the public good requires and so participate in government, without having to take responsibility for a view they are not fully confident in. Hence, the secret ballot is a *proper use* of anonymity, as it guarantees the equal participation of citizens in determining the character of public thinking.

Legislative representatives lie "downstream" of elections. They owe their position to citizens providing their best guess about the public good, and their job is to take those guesses and refine them, by deploying them in public deliberation and taking responsibility for public decisions made

on the basis of them. Since representatives cannot fulfill their duty of taking responsibility for public decisions if they voted anonymously, they are prohibited from doing so. Hence, anonymous voting by legislator is an *abuse* of anonymity, as it prevents constituents from appraising the thinking of their representatives.

Last is ballot initiatives – are they public or private? Should signatories to such initiatives be entitled to anonymity or must they sign in their own name? A clear answer is hard to discern since it seems that such signatories split the difference between citizens and legislators, acting a little bit like both. I argue however that signing a ballot initiative should be considered a public act since it is not part of an attempt to answer a question that is put to the people about the public good, but rather constitutes an attempt by one self-consciously organized subset of the public to put a question to the electorate, where the answer will not be subject to any further thinking, but will simply become law. Since ballot signatories attempt to query the electorate on schedule of their choosing, on a topic of their choosing, and without a chance for the electorate to rely on their representatives to answer it, they must assume the role of public spokesperson for it.

The organization of the chapter is as follows. In the next major section (section II) I illustrate the view that anonymity is not distinct and then argue that it is distinct by illustrating and defending the privileged participation view. I further illustrate the view by showing how it justifies anonymous speaking and reading. In the third major section (section III), I sketch a view of democracy in the course of discussing the three democratic practices listed above: voting in elections by citizens, legislative voting by representatives, and ballot initiatives. I argue that it is proper for the first to be anonymous, an abuse for the second to be anonymous, and the same for the last.

II. Theoretical Foundations

A. The Going View: Anonymity is Not Distinct

In this section I will offer some examples of the assumption that anonymity is only accidentally morally relevant, i.e., not distinct in itself. I do so to illustrate the nature of this assumption as well as its prevalence, but also to point out some of the insights that it has generated as well as to highlight some questions it has produced.

The view that anonymity is not distinct finds clear, albeit implicit expression in an article by Lyrissa Lidsky and Thomas Cotter.[14] Their focus is on anonymous speech and the general right to engage in it. They do not suppose that anonymity has a moral purpose that can be consulted to evaluate the appropriateness of such a right. Instead, they assume that such a right will be justified or not, depending on the effects that anonymity would have across speakers and speech situations in the aggregate, as well as the relative importance of those effects. Thus, anonymity for speech will call for a grand accounting. Costs and benefits are to be totaled up, and if the benefits sufficiently outweigh the costs to a greater degree than other options, anonymity is justified for speech, at least under circumstances as we find them.

As they acknowledge, such a grand accounting is difficult to carry out, because it depends on empirical facts that are hard to pin down. "In theory, audiences could be either better or worse off under a regime that grants strong protection to anonymous speech, as opposed to one that grants only weak protection, depending upon which effect – the production of more socially valuable speech, or the production of more harmful, though discounted, speech – predominates."[15] They assume though, as a kind of back-of-the-envelope estimate that persons are savvy and will approach

[14] *See* Lidsky & Cotter, *supra* note 6. This article is cited by other scholars focusing on privacy and anonymity. *See, e.g.*, Margot Kaminski, *Real Masks and Real Name Policies: Applying Anti-Mask Case Law to Anonymous Online Speech*, 23 FORDHAM INTELL. PROP. MEDIA & ENT. L. J. 815, 823 n.27 (2013) (citing and elaborating on Lidsky and Cotter).

[15] *Id.* at 1539-40.

anonymous speech with the correct amount of caution. They also assume that anonymity is a superior way of incentivizing speech, compared to other methods, such as greater institutional vigilance and consequences for retaliation against speakers.[16]

These empirical dependencies of their arguments are evidence of the assumption that anonymity is not distinct and should be evaluated by whatever effects it has in a situation. Put another way, anonymity is, on their view, the preferred way of calling forth a certain balance of good speech, but like any tool, it's only good given how things are now. Perhaps in the future, for all their argument claims, greater tolerance and protections against retaliation will make a right to anonymous speech unnecessary. The argument I give later will have the consequence that anonymity provides a distinct kind of protection for speech that has no substitute.[17]

A bigger problem though, and one that they also acknowledge, is the need for an evaluative framework. That is, even if we had rock solid empirics about how anonymity alters the content and quantity of certain kinds of speech, we still need to know how much bad anonymous speech can be tolerated in order to secure some amount of valuable speech that would not be spoken without the protection of anonymity. But this notion of a tradeoff is suspect, given that the constitutionally established value of the latter kind of speech is not usually balanced against the possibilities of the bad speech that might result.[18] For instance, the right to free speech itself is not dependent on there

[16] *Id.* at 1578. ("Indeed, for this class of speakers, a system that simultaneously compelled disclosure of authorial identity and effectively prevented retaliation would be preferable to one that merely protected anonymity, because (1) speech consumers would stand to benefit from knowing the speaker's identity, and (2) the speaker would stand a better chance of being taken seriously, all other things being equal. Reality suggests, however, that retaliation (let alone mere social ostracism) can never be prevented with 100% effectiveness, and thus that a rule forbidding anonymity almost certainly would discourage some apprehensive speakers from coming forward.").

[17] Though it's possible that circumstances will change such that very few persons feel the need to avail themselves of this kind of protection. My point in the text is that anonymity protects speech in a distinct way.

[18] Lidsky & Cotter, *supra* note 6, at 1581 n.190 ("Whether the costs and benefits of anonymous speech are even commensurable with respect to one another is debatable: As we suggested above, for example, if the autonomy interests in support of a right to speak anonymously are worthy of respect, how exactly does one determine the optimal tradeoff in return for a reduction in harmful speech? More importantly, and as others before us have noted, the social welfare approach appears inconsistent with a good deal of existing First Amendment jurisprudence…").

being a favorable balance of good speech to bad speech, but rather enjoys a justification that transcends this balance, precisely because the value of speech is to advance one's thinking, which happens through good and bad speech alike. Such are the problems of evaluating anonymity based on a grand accounting of various values, with no definite way to compare them.

Note that my goal is not to mount a full critique of their approach, but only to emphasize that it presumes a particular picture of what anonymity is and hence how it is justified. We are, on their approach, drawn into the project of creating a tradeoff schedule for the various effects of anonymity because it is assumed that anonymity does not provide its own, distinct moral logic, apart from totaling up costs and benefits. Recall that punishment has such a moral logic. One does not justify punishment by pointing to what one takes to be a sufficiently strong mix of good results or effects. Instead, certain facts are relevant to whether punishment is justified, because they ensure that the distinct good *of punishment* is realized in a particular situation.

Other scholars follow the approach of Lidsky and Cotter.[19] For example, Jeffrey Skopek mentions, at one point in a comprehensive and innovative article about anonymity, that it "often hides the identity of someone about whom facts are known for the purpose of putting such goods into public circulation."[20] This is, in my opinion, a correct characterization of anonymity that has profound implications. However, he does not develop this remark as a statement of anonymity's moral purpose or the key to its justification in particular cases. Instead, he intends it to be a summary of how anonymity (and attribution) "are both used by our law to shape the costs and benefits of creating goods in order to align private production incentives with public goals and

[19] Citations can be multiplied in law and philosophy. *E.g.,* Allison R. Hayward, *Bentham & Ballots: Tradeoffs Between Secrecy and Accountability in How We Vote*, 26 J.L. & POL. 39 (2010); Victoria Smith Ekstrand, *The Many Masks of Anon: Anonymity As Cultural Practice and Reflections in Case Law*, 18 J. TECH. L. & POL'Y 1 (2013) (cataloguing costs and benefits of anonymous speech); Julie Ponesse, *The Ties that Blind: Conceptualizing Anonymity*, 45 J. SOC. PHIL. 304 (2014) (noting the two-faced nature of anonymity).

[20] Jeffrey Skopek, *Anonymity, the Production of Goods, and Institutional Design*, 82 FORDHAM L. REV. 1751, 1755 (2014).

values, control information flows in order to address evaluation costs associated with using goods, and reallocate rights of control over goods in order achieve their efficient or fair allocation."[21] Anonymity is, for him, a peculiar or interesting kind of tool or "design lever" that institutions can manipulate to get the results they want, and will be justified when it is the right tool to bring about the desired outcome.[22]

Crucially though, Skopek explicitly disclaims any attempt to try and identify what the right outcome would be, such that the use of anonymity would be the most suitable lever to pull to make it so. From comments he makes throughout his article about balancing interests, it seems plausible that he is supposing that the right outcome is a complicated grand accounting of anonymity's impact on a multitude of goods like efficiency, freedom, and fairness, rather than a single, different good altogether; namely, the distinct good of anonymity.

The overall point is this. Many legal theorists of anonymity treat its justification as a matter of seeing to what degree it can call forth various mixtures of goods and bads in a given factual context. When it calls forth a sufficiently attractive mixture in a particular case, it is justified. This view of how anonymity is justified embodies a contestable assumption; namely, that we use it to realize goods, whatever they happen to be and as the opportunity arises, rather than as the designated means of realizing distinct moral goals. This is the assumption that anonymity is not distinct.

---

[21] *Id.* at 1756.

[22] Skopek at one point says that anonymity is not a "mere tool or aspect of privacy," id. at 1755, but it's important that this be taken in the right way. In saying this, he does not intend to deny that anonymity is a tool, but rather trying to claim that it is not a tool *of privacy*, but one of a wider sort. I will argue later that this is based on a misunderstanding of privacy, as many if not all the examples he cites of anonymity are cases of privacy, when privacy is understood properly as a certain kind of barrier to being held responsible.

B. Anonymity is Morally Distinct

In this section, I offer a theory of anonymity as namelessness, not unrecognizability.[23] The theory will address two issues: the nature of anonymity and its consequent role in moral life, with the second following from the first.

The nature of anonymity, I contend, is that of preventing persons from being held responsible for the things they do with ideas, e.g., speaking them and reading them. This aspect of anonymity's nature – that it is, I will say, *responsibility defeating* – is troubling and calls for justification. After all, it is a matter of general moral importance that persons be held responsible (fairly and proportionally of course) for those things that they *are* responsible for. For example, persons are responsible for how they move their bodies, and they may be held responsible for how they do so, through various means, such as civil suits for battery and negligence. There are exceptions, such that conduct that would ordinarily earn liability does not, but these exceptions, are, we say, privileges, and they require strong justification.[24]

Since persons are responsible for what they do with ideas, such as speaking them, it seems that anonymity, by preventing them from being held responsible for the same, grants an exception to the ordinary moral order of things, and so requires a strong justification as well. This justification is supplied by the proper use of anonymity, which is, to allow persons to participate in thinking with others, through externalizing and receiving ideas, without having first to work out which ideas they wish to be responsible for. This justification, it should be noted, is similar in some ways to what I believe to be the justification for the privacy of thoughts.[25] Persons *are responsible* for their thoughts,

---

[23] Margot Kaminski, *Real Masks and Real Name Policies: Applying Anti-Mask Case Law to Anonymous Online Speech*, 23 FORDHAM INTELL. PROP. MEDIA & ENT. L. J. 815 (2013) (discussing anti-mask statutes).

[24] 86 C.J.S. Torts § 28-34 (surveying some privileges against tort liability, such as the self-defense privilege and the litigation privilege).

[25] *See* Chapter 1 (arguing that thoughts are private because they prefigure but do not constitute what a person is willing to be held responsible for).

but may not be *held responsible* for them, since their thinking would suffer. Anonymity plays a similar role, but for communal rather than personal thinking. That is, anonymity allows a person to interact with the thinking of others, without having to "own" a particular position to do so. Still, I should stress, anonymity is not the ideal form of community thinking. The ideal is to think with others in one's own name, but getting to that point requires space in which to prepare and consolidate, and anonymity's job (and privacy's) is to provide that space.

Having laid out my two main contentions, I will now defend them. First is the claim that anonymity defeats the ability for others to hold one responsible. If one thinks that holding persons responsible involves imposing consequences on them or doing something to them, then anonymity will often preclude it, by denying the other party knowledge of the "target" for these interventions.

However, I argued in Chapter 1 that holding responsible does not necessarily involve imposing consequences on persons, though it might. Instead, I argued that the essence of holding someone responsible is the making of a mental record that links a particular person with something particular, e.g., an action or a thought, for which they are responsible. Congress holds agencies responsible by conducting "oversight" into their activities, journalists hold their subjects responsible by quoting them (putting them on the record), and friends hold each other responsible for their activities by taking account of them, often with more interest and precision than others would.

Because holding responsible is a matter of mentally linking a particular person with something particular for which they are responsible, the provision of these two pieces of information, the who and the what, in a combined form, is a way of asking or inviting another party to perform this bit of recordkeeping. This kind of asking or inviting constitutes *taking responsibility*. So, when person P speaks some content C *as hers* – "in her own name" – she provides, in a single package, everything that a listener needs to make the corresponding record: that P is responsible for

C. And her provision of all the requisites for being held responsible invites C to go through with it, and to hold her responsible. In an ordinary conversation, between one friend and another, this kind of exchange is ubiquitous. One person says that it's a nice day and the listener holds her responsible for saying so, by hearing and so recording her as saying it.

On this picture, one takes responsibility for some content by facilitating another's holding one responsible for it, by presenting it *as* one's own. The flip side is this: one avoids taking responsibility for some content by attempting to preclude another's holding one responsible for it, by *not presenting it as one's own*.

Crucially, there are two ways to not present content as one's own. One way is to *not present* the content at all. This was the subject of Chapter 1. If one does not speak one's thought to someone, then one *attempts* to keep them ignorant of its content (and the attempt will usually succeed, given that mind-reading powers and machines are not a part of our world), and through this attempt, signals an unwillingness to be held responsible for it. For the reasons given in Chapter 1, a person's unwillingness to be held responsible for a thought should, generally speaking, be honored.

However, a second way to not present content as one's own and so to signal an unwillingness to be held responsible for it is: to present the content, *but not as one's own*. Anonymity permits this second option. The reason is straightforward. Holding responsible is, I argued, a matter of attributing something to someone as theirs – of mentally linking a specific person to some specific content – but without access to the person through their name, this is impossible. If holding responsible is a matter of recording something in one's mental "file" on a person, then without a name, the right file cannot be updated. For example, if one receives an anonymous note, one can, at

most, make a mental note of the content, and file it under the heading "?", as a placeholder for whoever sent it.

This theoretical result is bolstered by common intuitions about names. When something bad happens and we want to hold someone responsible, we may shout "I want names!" Informants are threatening because they might "name names," and when we want to praise or criticize someone specifically, we may "call them out" by name. In all of these cases, a name provides the key for holding someone responsible, because it provides a label, shared among a community, under which their actions can be recorded, and, if necessary (and it often is), judged in various ways.[26] Put simply, a name is the mental means of record-making about a particular thing, *par excellence*.[27]

An example of these claims is art. An artist signing a painting is like a speaker speaking a thought in her own name. The artist's signature is not just a piece of evidence that she painted it, on a par with a chemical or brushstroke analysis that says the same. Rather, the artist's signing a painting is an endorsement. It indicates that she thinks it counts as a full member of her corpus. A signed painting is one she wishes to be known or evaluated by, whereas an unsigned painting is one she disclaims, as not a full representative of her artistic style. Again, the point is not that disclaiming can make it so – she painted the painting. Nothing can change *that*. Rather the point is that any interpretation of her art should take into account her view of her own work. All the paintings that she painted cannot be lumped together as her corpus, because there are distinction between them in terms of how she regarded them.

[26] The community is the holder of the "complete" record about the person, in the sense that whatever others attribute to her under her name (things are complicated for Batman, or others with "secret identities") can, in principle, be combined, e.g., "Oh you know Jim? I was on a trip with him once and he …."

[27] *Cf.* Jeffrey Skopek, *Reasonable Expectations of Anonymity*, 101 VA. L. REV. 691, 720-22 (2015) (arguing that a name is just like any other piece of information in being a potential way to single the person out such that "namelessness is neither a sufficient nor a necessary condition for anonymity.").

The discussion so far has addressed the first goal of this section: to elaborate the nature of anonymity, which is that it is responsibility defeating. Since holding responsible is, at its most basic, an act of mental recordation; of linking, in thought, a particular act or thought with a particular person, and since a name is the way one thinks directly about a particular person, ignorance of person's name precludes holding them responsible. If one does not know who wrote a note, its content cannot be attributed to anyone as that for which they are responsible. It is disclaimed content.

It bears emphasis that though, as I just said, namelessness precludes holding responsible, it is not true that anonymity guarantees namelessness. Not signing a note or a painting may often keep it nameless, but not always and perhaps not forever. The author may be easy to identify despite there being no signature, or it may come to light later, when new information is revealed. Still, and this is what I wish to emphasize, the anonymous speech is still a *bid* or an *attempt* to avoid being held responsible for content, and this aspect of anonymous speech is inseparable from it and survives the identification of its speaker. What I mean is that if the anonymous author of a note is later identified, it is not as if the content of the note can now be considered written in the author's name, in the same way as her other correspondences.[28] No, a cloud of uncertainty or ambiguity hangs over the note, even after being linked to its source, because the linkage is ersatz. The note was never put forward *as* the author's, and this fact is not altered by finding out that she did in fact write it. Instead, its disclaimed status lives on.

Taking into account these last points, my claims about the nature of anonymity can be put compactly like this: a nameless person cannot be held fully responsible for the content she puts

[28] The same is true when an anonymous painting is later identified as the work of a particular artist. Though she painted it, the import of the painting to interpretations and judgment of her artistic legacy is ambiguous due to her judgment that it was not a full member of her corpus.

forward, and anonymity always seeks namelessness, even if it fails to achieve it. This summary allows an equally compact statement of the parallel I see between punishment and anonymity. Punishment is for holding persons responsible for their crimes, and anonymity is for preventing speakers from being held responsible for their speech.

Having dealt with the nature of anonymity, I turn now to its value. What's the good in it? Why should persons be allowed to defeat the responsibility for which ideas they traffic with? Why should they have this privilege? My answer is built around thinking, in a way that recalls my arguments from Chapter 1. I briefly recount the argument in order to show how it applies to anonymity.

The essence of my argument in Chapter 1 was that persons have an interest in thinking well, and that thinking goes well when persons have control over when their solitary thinking becomes a part of thinking with others. Without this control, persons may be misrepresented or have their thinking stunted or diverted. With this control, persons can wait until they are satisfied with their thinking before integrating it with the thinking of others, as the latter requires carrying out a demanding joint activity and fulfilling its attendant obligations.

Morality recognizes this interest in control by prohibiting others from accessing the content of thoughts that one has not signaled a readiness to reveal. This prohibition against access protects one's ability to withhold content from joint-thinking and to introduce it to others only selectively, at one's own pace and on one's own terms. It allows persons to start thinking by themselves and then pull in mentors, friends, and spiritual advisors before broadening to, say, a small workshop, and then to a conference, and perhaps, eventually, to the republic of letters, in the form of published work.

Protection for growing the breadth of one's own thinking at one's own pace is valuable, but it's also limited, because persons also need access to others *to* grow their thinking. Moreover, one

sometimes has an interest in preparing to think with an audience by thinking, preliminarily, with that very audience, and not a different, proxy audience with which one is more comfortable with. But if one were only protected from the thinking of others so long as one kept to oneself, then one would, to gain from joint thinking with a particular audience, always have to take responsibility for the relevant thought in front of that audience. This is a catch-22, because in order to think with some others about a thought, so as to better assess whether to take responsibility for it, one would first have to take responsibility for it.

These points illustrate the need for protection from the thinking of others that is consistent with participating in various community discourses rather than withholding one's thoughts from them.

Take thinking with the public at large. Public thinking is not just private thinking with a lot of people. It's of a different order. For example, publishing something does not just make it available to a certain set of persons, but to persons as such – to beings who traffic in ideas.[29] Correspondingly, participation in public thinking is distinctly valuable – both in itself and as a means of improving what one thinks[30] – and also particularly intimidating because it defines who one is, not just in some circles, but to thinkers as such. Thus, persons have, as regards the public, a need for a participatory kind of protection against being held responsible for their thoughts – such that they can think publicly without already having a position that they are ready to take responsibility for.

Anonymity is this kind of protection. Because it prevents persons from being held responsible for their thoughts by suppressing their identity and not their ideas, it allows persons who are unready to be held responsible for their ideas to participate, to an extent, in public discourse. For

[29] *Cf.* Michael Warner, *Publics and Counterpublics*, 14 PUB. CULTURE 49 (2002) (discussing the formation of a public).

[30] In thinking with others, one must articulate ideas in ways that can speak to many different kinds to listeners, and likewise to respond and address arguments from many different kinds of challengers or critics.

this reason, anonymity can be understood as the public-facing or participatory side of mental privacy. This is its moral logic: to permit persons to engage in community thinking (with the maximal community being the public) without the burden of having to take responsibility for their thoughts.

To conclude, I will show how this moral logic works to justify the general First Amendment rights to read and speak anonymously. Considering these core cases will draw out the key claims of this section.

Consider anonymous reading first, which includes anonymously preparing to read, say, by querying a search engine.[31] It is a paradigmatic illustration of the moral logic of anonymity: the thinking-based, participatory interests persons have in reading are enormous, the risks posed by reading in one's own name are serious, and the possibilities for abusing anonymous reading are scarce.[32]

First, the interest in receiving the content of others is crucial to thinking, and without it, the ability to have thoughts without revealing them to others would be of limited value, since one would be restricted to whatever thoughts one could put together by oneself. One needs material to gain from solitary thinking. One needs to access the rich legacy of past thinking in the form of "great"

[31] The First Amendment right to read anonymously has not received an explicit endorsement from the Supreme Court. Nonetheless, *Stanley v. Georgia* strongly supports such a right, since it permitted persons to view obscene material in the privacy of their home. Justice Marshall conceived of the case in terms of the privacy of the mind, and characterized the right at issue as "the right to be free from state inquiry into the contents of his library." Stanley v. Georgia, 394 U.S. 557, 565 (1969).

[32] Others have also argued forcefully that anonymous reading is important. *See, e.g.*, Marc Blitz, *Constitutional Safeguards for Silent Experiments in Living: Libraries, the Right to Read, and a First Amendment Theory for an Unaccompanied Right to Receive Information*, 74 UMKC L. REV. 799 (2004); Marc Blitz, *The Freedom of 3d Thought: the First Amendment in Virtual Reality*, 30 CARDOZO L. REV. 1141 (2008); Julie Cohen, *A Right to Read Anonymously: A Closer Look at "Copyright Management" in Cyberspace*, 28 CONN. L. REV. 981 (1996); Neil Richards, *The Perils of Social Reading*, 101 GEORGETOWN L. REV. 689 (2013).

books, but also to get in touch with the zeitgeist, which may live in pulp novels or popular TV. Seeing what others are thinking is a way of participating in community thinking at a basic level.

These benefits are threatened if one must read in one's own name, because what one reads is an infirm but tempting basis for others to reach conclusions about who one is. Consider two examples, one old and one new. In the late 80s, a reporter gathered Supreme Court nominee Robert Bork's video rental records from his local store.[33] Nothing too shocking was revealed, but still, this exposed, inchoately, the workings of Bork's mind. Much more recently, the *New York Times* ran an article that analyzed celebrity bookshelves (made visible due to zoom calls prompted by the coronavirus pandemic) in an effort to see what they "revealed" about them.[34] Here too, the content a person engages with excites speculation about who they are and what they think. Reading on the subway also testifies to the dynamics of judgment at play in such knowledge.[35] One may be circumspect about reading notorious works. No doubt, what one reads reflects what one thinks, but only inchoately, because reading is how one explores an area, or satisfies curiosity, in preparation for settling on what one truly believes. Search queries are an even earlier stage of one's curiosity, in which one decides *what* to read.

Moreover, it is often the most "radical" material that is crucial to developing one's own views, because they help one the chart the boundary of what is ambitious and extreme and what is

---

[33] Andrea Peterson, *How Washington's Last Remaining Video Rental Store Changed the Course of Privacy Law*, WASHINGTON POST, April 28, 2014 (https://www.washingtonpost.com/news/the-switch/wp/2014/04/28/how-washingtons-last-remaining-video-rental-store-changed-the-course-of-privacy-law/).

[34] Gal Beckerman, *What Do Famous People's Bookshelves Reveal?,* N. Y. TIMES, Updated July 27, 2020 (https://www.nytimes.com/2020/04/30/books/celebrity-bookshelves-tv-coronavirus.html).

[35] Susan Dominus, *Snoopers on Subway, Beware Digital Books,* N. Y. TIMES, Mar 31 2008 (https://www.nytimes.com/2008/03/31/nyregion/31bigcity.html) (discussing privacy in what persons read in public); Scott Rogowsky, *Taking Fake Book Covers on the Subway* (https://www.youtube.com/watch?v=jFxu9dOO4zk&feature=youtu.be) (humorously dramatizing the privacy aspect of reading on public transportation by conspicuously reading fake books with outlandish titles).

nonsense or patently mistaken.[36] It's a sensitive affair, because in reading, even when one is reading mainstream ideas, one plays host to another mind, and gives oneself over to its narrative of the relevant issues and ideas, or, in a novel, to its plot and characters.[37] In reading an author, one becomes her associate or fellow traveler for a time, and the risk of "guilt by association" is great (in the same way that literal association with an organization or a person may be tainting, even though one's associates hardly define one's own viewpoint).[38] A prominent example is reading communist literature during the fifties and sixties.[39] Going much further back, we can take account of the decisive change in society that resulted from the slow introduction of silent reading starting in the 16th century.[40] Reading something without nearby others knowing what it was "radically transformed intellectual work."[41]

Finally, anonymous reading can be abused; to infringe copyright[42] or to learn information for malevolent purposes,[43] or to serve as the audience for wrongful revelations about others, such as by watching revenge porn. These consequences may be managed in other ways though.[44] There are ways to protect copyright besides trying to reveal the identity of those who read anonymously or by blocking anonymous access, and gaining information about how to break the law does not entail a

---

[36] It's a separate issue, but one worth mentioning, that anonymity is, in this regard, also crucial for those propounding minority viewpoints, because if the latter cannot even be considered without incurring social judgment, then they will be robbed of their persuasive merit before any discussion takes place. *See* Margot Kaminski and Shane Witnov, *The Conforming Effect: First Amendment Implications of Surveillance Beyond Chilling Speech*, 49 U. RICH. L. REV. 465, 508 (2015) (making this point).

[37] CECILE JAGODZINSKI, PRIVACY AND PRINT (1999) (discussing religious authorities' worries about the effect of reading on devotees); Georges Poulet, *Phenomenology of Reading*, 1 NEW LITERARY HISTORY 53 (1969);

[38] NAACP v. Alabama Ex Rel. Patterson, 357 U.S. 449 (1958).

[39] *E.g.,* Keyishian v. Bd. of Regents, 385 U.S. 589 (1967) (First Amendment bars requirement to certify one is not a communist).

[40] 3 Roger Chartier, *The Practical Impact of Writing, in* A HISTORY OF PRIVATE LIFE 124 (Roger Chartier ed., Arthur Goldhammer trans., 1989).

[41] *Id.* at 125.

[42] *See* Cohen, *supra* note 32.

[43] *See* Eugene Volokh, *Crime-Facilitating Speech*, 57 STAN. L. REV. 1095 (2005) (many of the possible benefits of such speech redound to readers of it, some of whom, perhaps most, will be anonymous).

[44] *Cf.* Blitz, *supra* note 32, at 1171 ("a silent listener threatens less disruption to the lives of others than does a public speaker.").

desire or ability to do so much more any non-mental step in doing so. There is ample time to intercept criminal activity without trying to monitor what people are wondering about.[45] The same is arguable true with anonymous access to privacy-invading content, as those who post such information may be better targets for identification, though there may be cases in which a more comprehensive legal strategy is warranted.

Anonymous speaking is also highly justified, though its abuse may directly constitute a violation of a non-property law, such as in the case of anonymous defamation, or the interference with the thinking of others through the spread of false information. Nonetheless, speaking anonymously is an important channel for participating in public reasoning. It allows one to make an idea a part of public discourse. It also allows persons to test out ideas by soliciting responses indiscriminately from others, and in turn, to identify those who are interested and knowledgeable. These justifications for anonymous speech do not depend on the need to resist or obviate the wrongful conduct of others.

In other cases though, one may need anonymous speech to (partially) counter the wrongs of others. For instance, someone with an unpopular viewpoint may be perfectly satisfied with it, but wish to speak it anonymously for fear that they will unfairly suffer for it, whether due to reprisals, or else due to being stereotyped by others. A whistleblower may be confident that wrongdoing has occurred, but still resist speaking without anonymity, due to the risk of reprisals. A member of a radical political party on the other hand may be worried about being misrepresented or demonized. The Supreme Court has paid special attention to the interests of political speakers in remaining

[45] In Chapter 3, I argue in favor of the thought crime doctrine, which prohibits the punishment of beliefs. If that argument is sound, it lends support to the claim that they should be able to read anonymously, as reading something does not even, necessarily, form a belief, and if persons are permitted to harbor bad beliefs without others knowing, it seems they have an interest in being able to see if they wish to harbor bad beliefs without others knowing.

anonymous, arguing, in an important precedent, that "identification of the author against her will is particularly intrusive; it reveals unmistakably the content of her thoughts on a controversial issue."[46]

In still other cases, one does not wish to take responsibility for an idea, not because one is uncertain of it, nor because one will suffer for it, but rather because one wants the idea to be fairly received, and appreciated for the true value it has. An example is art. A novelist may be satisfied with her novel and satisfied that she can publish it without suffering any reprisals. Nonetheless, she may suspect that because she is a woman, her novel will be poorly received. This mis-reception is an indirect misrepresentation of her, since the neglect of her novel will reflect badly on her as a novelist. It may also be injurious to her directly, as the neglect of her art itself, which she is invested in.

To avoid these responses, she may publish under a pseudonym or anonymously. In this example, the author is subject to wrongful prejudice, but in other cases, the prejudice may not be wrongful, at least not in the same way. For instance, J.K. Rowling published under a pseudonym to receive more honest feedback from editors, and many presidents have done the same.[47] Here the goal is not to avoid prejudice but to avoid superficial praise, earned only by their reputation rather than their ideas. Publius is another famous example from American history. Though it is a pseudonym, it wears its pseudonymity on its sleeve, rather than purporting to be the true name of someone.

Speaking anonymously can be abused. This is because anonymity is a departure from the norm, and the reason for the departure cannot often be easily discerned without the aid of

---

[46] McIntyre v. Ohio Elections Comm'n 514 U.S. 334, 355 (1995). *See also*, Talley v. California, 362 U.S. 60 (1960) (anonymous political pamphleting); Buckley v. Am. Constitutional Law Found., Inc., 525 U.S. 182 (1999) (petition circulators could not be made to wear name tags); Watchtower Bible & Tract Soc. of NY, Inc. v. Village of Stratton, 536 U.S. 150, 166-67 (2002) (door to door religious solicitation).

[47] Sylvia Hui, *J.K. Rowling Revealed as Writer of Crime Novel*, WASH. POST (July 14, 2013), https://wapo.st/2UEpxBm.

background conventions or institutions. Consequently, some anonymous statements, if they are false or abusive, or even if they originate in tentativeness on the speaker's part, may not be communicated, because others may take the anonymity as a sign that what is being said *is* true, but told anonymously because it is threatening to someone in power. An example is an anonymous tell-all book, say, from an "insider" in a presidential administration.[48] In such case, readers will infer that anonymity is being used to avoid wrongful reprisals, which provides cover for someone to spread disinformation more easily. Denials of the falsehoods may fan the flames too, since they may appear to be evidence of a cover up, and proof of the damaging nature of the revealed secrets. In other cases, such as the anonymous posting of non-consensual pornography, anonymity is nothing but an attempt to avoid being held responsible.

It is important to see the significance of these abuses. A legal right to speak anonymously may incidentally shield, to a degree, abuses of that right, given that the proper use of the right is responsibility defeating. Due process may require additional steps to hold person responsible for anonymous defamation, simply because without such steps, the legitimate value of anonymity would be undermined. However, I think there is little doubt that using anonymity simply to escape being held responsible for wrongdoing is an abuse of it. If so, this is further evidence that anonymity has a moral logic. Anonymous speech that is otherwise illegal should not be evaluated on a cost-benefit basis, with some credit being given to it simply because it is a way of organizing one's message. The

[48] *E.g.,* ANONYMOUS, A WARNING (2019) (insider look at the Trump White House). The author of *A Warning* was later revealed to be Miles Taylor, who defended his anonymity in terms of the impact he intended his book to have on public discourse. Katherine Faulders, *'Anonymous,' Author of White House tell-all book, revealed to be Miles Taylor*, ABC NEWS (Oct. 28, 2020), https://abcnews.go.com/Politics/anonymous-author-white-house-book-revealed-miles-taylor/story?id=73884296 ("Issuing my critiques without attribution forced the President to answer them directly on their merits or not at all, rather than creating distractions through petty insults and name-calling. I wanted the attention to be on the arguments themselves.").

anonymity in such a case is no different than waiting until dark to commit a crime – the goal is to escape detection for what is wrongful.

For the above reasons, persons have a strong interest in being able to interact anonymously with communal thinking, with anonymous speaking and reading being two complementary sides of that activity. These two kinds of interactions with the republic of letters – the public understood as the public sphere of speech – are legally protected by the U.S. constitutional system in the form of the First Amendment. In the next section, I explore whether and to what degree persons have any interests in interacting anonymously with the republic itself – the public understood as a constitutional order.

III. Democratic Applications of the Theory – Citizen Voting, Legislator Voting, and Citizens as Legislators

Should democracy always be practiced in the open? To investigate this question, take, as a focal trio, voting by citizens,[49] voting by legislative representatives,[50] and signing a ballot initiative, which amounts to a case of citizens attempting to legislate by proposing a law directly to the electorate for approval or rejection.[51] The issue at the heart of these practices is: to what degree must political power be exercised openly, and what are the factors that bear on the answer to that question in a particular case? A more focused version of this question is: what is the nature of a citizen's interest in voting anonymously? Is the secret ballot for citizens a pragmatic attempt to

[49] *See* Levmore, *supra* note 11, at 2219-2220 (pointing out the prevalence of secret balloting in national elections). An exception is West Virginia. *See* W. Va. Code Ann. § 3-1-4 (West) ("In all elections the mode of voting shall be by ballot, but the voter shall be left free to vote by either open, sealed, or secret ballot, as he may elect.").

[50] Some theorists have wondered why elected representatives must vote publicly. *See* Levmore, *supra* note 11, at 2219-2220 ("Anonymity is thus required rather than simply accepted in (the legal rules governing) general elections. In contrast, when voting takes place in representative or deliberative bodies, it is open."); Annabelle Lever, *Privacy and Democracy: What the Secret Ballot Reveals*, 11 LAW, CULTURE, AND THE HUMAN. 164, 175 (2015) ("While democratic legislators may be more vulnerable to intimidation than citizens – as they are relatively few in number, and hold special power and authority *qua* legislators – it is the former, not the latter, who must vote openly, not secretly.").

[51] *See* John Doe v. Reed, 561 U.S. 186 (2010) (Oregon ballot initiative); Buckley v. Am. Constitutional Law Found., Inc., 525 U.S. 182 (1999) (analyzing Colorado ballot initiative rules).

prevent corruption and intimidation or is it a more fundamental interest, akin to the right to speak anonymously or the right to participate in politics?

The Supreme Court has never directly taken up this question, but it came close in *Doe v. Reed*, in which it considered a ballot initiative to repeal a recently passed Oregon law altering the rights of domestic partners, including same-sex partners.[52] According to state law, the signatories to the ballot initiative and their addresses were public record and could be requested by private parties under a state transparency statute. The law was facially challenged as a violation of First Amendment in virtue of requiring the disclosure of a persons' political affiliations. The decision was 8-1 in favor of the holding that a state's choice to make ballot initiative signatories publicly available was permissible under the First Amendment. However, several justices noted that the challengers to the law could prevent disclosure if they could show a specific risk of reprisals or harassment from publication. From the opinions, three broad views on the secret ballot can be extracted, though that issue was not explicitly discussed.

One view, embodied by Justice Thomas in dissent, sees "privacy in political association"[53] as vitally important, and so allows him to dispose of the case easily. State law may not require that ballot initiative signatories be made public because it chills the exercise of First Amendment electoral rights.[54] Given the privacy justification he offers for this claim, it seems clear that he would support the secret ballot as earning significant constitutional support.

In his concurrence, Justice Scalia takes the opposite view, which is that the secret ballot enjoys no fundamental constitutional or democratic justification. He argues this by invoking the

[52] John Doe v. Reed, 561 U.S. 186 (2010).
[53] *Id.* at 231 (Thomas, J., dissenting).
[54] *Id.* at 229 (Thomas, J., dissenting).

history of electoral practices in the U.S.,[55] but his more interesting theoretical argument is one that operates by invoking a principle, according to which acts of legislative effect must be public. He argues that the "public nature of federal lawmaking is constitutionally required," and he clearly intends this constitutional judgment to encompass state lawmaking too.[56]

This principle disposes of the case, given his opinion that, by signing a ballot initiative, a signatory is "acting as a legislator."[57] This principle is broad though, and also arguably entails that there is no First Amendment right, or fundamental democratic interest, in voting anonymously – something that Chief Justice Roberts correctly points out in a footnote.[58] Scalia's view can be characterized as consistently in favor of publicity when legislators vote, when citizens vote as something akin to legislators, and when citizens vote. This view puts him in the company of many political scientists and theorists, who argue that, *prima facie*, voting by citizens should be public, but that sometimes the dangers of voter corruption (e.g., vote buying) or voter intimidation (threats) warrant secret balloting as a prophylactic counter.[59]

All the other justices are in the middle, in that they gravitate toward either Justice Scalia's or Justice Thomas' position about the fundamentals of the case – about whether the right staring point is privacy of political belief or the publicity of electoral activity – but find reason to moderate it for

---

[55] *Id.* at 221 (Scalia, J., concurring) ("Our Nation's longstanding traditions of legislating and voting in public refute the claim that the First Amendment accords a right to anonymity in the performance of an act with governmental effect.").

[56] *Id.* at 222 (Scalia, J., concurring).

[57] *Id.* at 221.

[58] *Id.* at 196 n.1.

[59] Most famously, John Stuart Mill. JOHN STUART MILL, CONSIDERATIONS ON REPRESENTATIVE GOVERNMENT Ch. X (1873) (available at https://www.google.com/books/edition/Considerations_on_Representative_Governm/16FFAQAAMAAJ?hl=en&gbpv=0). Contemporary examples are Daniel Sturgis, *Is Voting a Private Matter?* 36 J. OF SOC. PHIL. 18 (2005); Philip Pettit & Geoffrey Brennan, *Unveiling the Vote*, 20 BRIT. J. OF POL. SCIENCE 311 (1990); Pierre-Etienne Vandamme, *Voting Secrecy and the Right to Justification*, 25 CONSTELLATIONS 388 (2017) (arguing that balloting itself should be secret, but persons should have to give their justifications for voting as they did). *See generally* Tom Theuns, *Jeremy Bentham, John Stuart Mill and the Secret Ballot: Insights from Nineteenth Century Democratic Theory*, 63 AUSTRALIAN J. OF POL. & HIST. 493 (2017) (offering a good, historically informed view of the issue).

pragmatic reasons, from either direction. For instance, Justice Roberts essentially endorses Justice Thomas' view that anonymous political participation is a fundamental First Amendment right, but unlike Thomas, he sees the government interests in protecting against fraud and mistakes as enough to justify publicity despite that right. For him, ballot initiatives implicate a fundamental privacy interest, but that interest may be qualified when the pragmatics of the situation demand it.[60]

Justices Sotomayor, Ginsburg, and Stevens on the other hand come at things from Scalia's angle, favorably quoting his views about the publicity of lawmaking, and lauding, in broad terms, the importance of "openness" and "transparency" as well as the "inherently public" nature of ballot initiatives. However, exactly how their position relates to his is unclear. On the one hand, they do not rest their opinion solely on the public nature of legislating, but instead argue, in addition, that pragmatic interests permit the regulation of elections by states and that any free speech interest in anonymous political association is only minimally implicated by disclosure of the ballot initiative signatories. On the other hand, they argue this last point by analogy to campaign finance disclosure rules. These rules, they say, "do not prevent anyone from speaking."[61] But public voting would not prevent anyone from speaking (or voting) either. And so, if this latter argument about the negligible effects of disclosure is given precedence, then their position is closer to Scalia's in favor of thoroughgoing publicity. If the pragmatic rationales are more important, then it seems that they are closer to Justice Roberts in believing that there is a privacy interest in political association that can be curtailed for pragmatic reasons.

The lesson I want to take away from *Doe* is about its dialectic. The two principles that divide the justices are blunt and so they gobble up the intermediate moral terrain, with only pragmatic concerns tempering their reach. Political beliefs are either private or else voting must be public. I

[60] *Id.* at 197-99.
[61] John Doe v. Reed, 561 U.S. 186, 196 (2010).

suspect the bluntness owes to the fact that both principles are likely to be the consequences of deeper, more tangled moral principles, as evidenced by the fact that it seems coherent to ask why lawmaking should be public, and why a person's political beliefs may be subject to privacy protections. If we knew the answers to these questions, we might be able to see how there is room for distinctions between democratic practices that do not amount to pragmatic limitations on a single, broad principle. Put simply, there may be a fundamental interest in some anonymous electoral action *and* a fundamental need for public lawmaking.

I try to make the case for these claims by arguing that many institutions in a democracy require the participation of the people, and that the kind of participation that they require is a form of thinking, taking the form of an answer to a question. I further argue that the publicity or privacy of their response is determined by who should take responsibility for it, which is in turn a matter of the role that the question and its answer play in the thinking of the polity. Since different institutions answer different questions, and at different moments in a polity's attempt to grapple with them, the privacy or publicity of various electoral processes is not something that is determined by a single principle, but by the dimensions of the thinking that needs to be done.

A good introduction to this way of thinking is by considering the relationship of a grand jury to a standard trial jury, usually referred to as a petit jury.[62] Both are institutions of democratic participation. They are both places in which "the people" are consulted, though of course, not in its entirety, but in the form of a representative cross-section of ordinary citizens. We could ask the general question of whether citizens should have to take on the role of juror (of any kind) in their own name or if they may do so anonymously, but I think it's fairly evident that this would be a

[62] Civil petit jurors more regularly consider "mixed" questions of fact and law, and I think this earns some justification when civil law is mainly precedential and so not systemized by legislative activity. Owing to this fact, a civil jury is given a wider ambit to consider the propriety of rules that are, largely, court designed.

cramped way of putting the question and would prefigure a cramped answer, given that the deliberation of both of these bodies is starkly different.

A grand jury and a petit jury are both entitled to some measure of secrecy in their deliberations, with the grand jury members being anonymous by default, whereas the members of a petit jury are not anonymous by default, but can be granted anonymity if the situation warrants it.[63]

The secrecy of the deliberations of both grand and petit juries is justified by the need to receive the public's unvarnished opinion.[64] However, the anonymity of a grand jury owes, I claim, to the role it plays in the stretch of government action that aims to convict someone for a crime. Through a grand jury indictment, a prosecutor becomes more than just a government employee; she earns the backing of the public, such that when the case actually comes to trial, the prosecutor's side can rightly be titled "the people."

In this way, the grand jury is preliminary kind of thing. It is to the people what a thought is to a person. What a person thinks is the content she is mulling over, in preparation to make it public. Similarly, a grand jury indictment is a public prosecution that is in the works. It is what the state thinks happened, though it has not yet said so. The saying so will happen when the accusation is made and the indictment is acted on. It should be noted that the kind of preparation that the privacy of the grand jury permits is applicable to persons. Individuals too may wish to prepare for taking a position by gathering and organizing documents and evidence that will support or explain

[63] *See* 2 Fed. Grand Jury § 16:2 (2d ed.) (principle of secrecy including anonymity); United States v. Mansoori, 304 F.3d 635, 649 (7th Cir. 2002) (discussing when an anonymous jury can be empaneled).

[64] Roger A. Fairfax, Jr., *Grand Jury Discretion and Constitutional Design*, 93 CORNELL L. REV. 703, 748 (2008) ("Secrecy also shields the grand jury's exercise of discretion from public glare, thereby minimizing the possibility that grand jury members will feel compelled to base their decisions on concerns about immediate public backlash in a given case. Thus, secrecy can lead to greater reflection and richer, more sincere deliberation.").

what they say.[65] Having this support ready may make their claims more persuasive, but also more demonstrably considered.

Another way of making my point about a grand jury's preliminary status is through the idea of sincerity. When a person says what she believes, she is sincere, and when a prosecution charges in line with an indictment, the effect is similar. The government displays good faith or earnestness, given that it does not rely solely on its own judgment of what has happened, but has recruited another set of "independent" minds for confirmation.[66]

A petit jury, by contrast, is to the people what speech is to a person. A petit jury is not preliminary or preparatory. Instead, it is supposed to publicly deliver a verdict and to actually decide whether someone has committed a crime or not.

These differences in their roles lead to differing needs for anonymity. A grand jury does a more independent, open-ended kind of thinking than a petit jury, and hence needs more independence from the thinking of others. The grand jury is empowered to make its own way through the evidence. It can make decisions about what witnesses to call and whether to expand its inquiry. It can make tough tradeoffs between which persons to pursue for which crimes. In brief, the grand jury is empowered, to an extent, to shape the question that it answers.

By contrast, a petit jury (in a criminal trial) is asked to answer a specific factual question about what happened. It does not make tradeoffs and is not able to control the evidence that it comes into contact with. Instead, evidence rules dictate what the jury can be exposed to and adversary procedure dictates what arguments the jury here. The jury is thus in a better position to take responsibility for its thinking. Its question is narrow, it is exposed to the fully prepared

[65] Thanks to Barbara Herman for bringing this point to my attention.

[66] Search warrants are similar in that the executive recruits the judiciary as confirmatory source of thinking. The judiciary plays the role of an "independent" here.

presentations of both sides of the accusation, and the evidence it receives is filtered through rules of evidence.

A final difference is the effect that a grand jury indictment and a petit jury verdict have. The former is a serious matter – it's an accusation – but it is also a starting point rather than an ending point. The indictment is the gateway to the most robust aspects of an adversarial system. A jury verdict however is a crucial ending point, in that it may large a role in restricting someone's liberty. It thus must be carried out in the open. Citizens have a great interest in the decision by which a person's defense of their innocence is possibly rejected in favor of the state's narrative about what happened and what is to be done about it. The preceding points about the jury's question bolsters this point. Ordinary citizens can follow along with a criminal trial and see the evidence and assess it.

The point of this discussion is to show that the publicity or anonymity of citizens in the performance of democratic obligations depends on the justifications for those obligations and the setting in which they must be performed. In a (federal) criminal trial, the public is present in different roles. The prosecutor represents the people by acting on their indictment, the judge represents the people by being selected by them, (usually) through their representatives, for public office, and the jury represents the people by being a fair cross-section of them. These different roles combine to produce a display of public reason, which amounts to the accumulated effect of different persons thinking about a subject matter from different angles, organized in a productive way, that comes to a conclusion in an adversarial trial. Grand juries set the stage for thinking by ensuring that the prosecutor is proceeding in good faith. Petit juries bring things to a close by making a judgment with more or less decisive effect, after all the evidence is in.

I now turn to discussing non-judicial democratic practices: voting by representatives, voting by citizens and then ballot initiatives. These differ in significant ways from jury service, but they are

similar in that they involve the people addressing a question, where their answer will be part of a stream of thinking involving different persons and perspectives, designed to produce a strong instance of public thinking.

1. Representatives Must Vote Publicly

I will start with a view of representative democracy according to which it is primarily concerned with, and partially justified by, the creation of public persons. Celebrities are public persons, but they are so *de facto*. By contrast, a representative democracy creates public person *de jure*, for the purpose of executing and implementing public business. The crucial role that representatives play is not that they make decisions with binding legal effect, though of course this is important. Rather, the valuable role that public persons play is one that is indispensably played by persons, which is to take responsibility for thinking. Representatives take responsibility for public thinking in the specific sense that they heighten and develop it, by combining the disparate viewpoints of their constituents into what is ideally, a coherent, reasoned, and accessible vision of the political landscape.[67]

Though I will not here defend a full theory, I mean to disclaim views according to which representatives are just the agents of the people, and so may not do more than carry out their instructions. Instead, I think the relationship between the people and their representatives is more reciprocal and open-ended, with representatives taking account of what their constituents think but also reflecting their views back to them, with some added polish. A simple example might be John McCain telling a town hall participant that Obama was not an "Arab." I do not mean to suggest that

[67] THE FEDERALIST NO. 10, at 44 (James Madison) (Terence Ball ed., 2003) ("The effect of the first difference is, on the one hand, to refine and enlarge the public views, by passing them through the medium of a chosen body of citizens, whose wisdom may best discern the true interest of their country, and whose patriotism and love of justice will be least likely to sacrifice it to temporary or partial considerations. Under such a regulation, it may well happen that the public voice, pronounced by the representatives of the people, will be more consonant to the public good than if pronounced by the people themselves, convened for the purpose.").

McCain's response was flawless, but only that it illustrates the way in which information flows in two directions between constituents and those they elect.[68]

I want to elaborate this point briefly in terms of the question answering function of democratic actions. Above, I discussed the way in which a petit jury represents the public in an orchestrated bit of public thinking that is the trial. Elected legislative representatives are similar. They are the representatives of the public in the orchestrated bit of public thinking that consists, at least, in legislating: making speeches and voting in Congress. In other capacities, the role of the legislator is different, but still derived from this public function. For instance, legislators may have meetings with constituents or lobbyists. Such meetings are not themselves legislating, but are also not private in the way that a legislator may have a friend over for dinner, and the discussion turns to politics. The lobbying may be rightly subject to duties of, for example, disclosure, given that it is thinking that is pursuant to legislating.[69]

Unlike jurors, representatives are elected and not randomly selected, and this is significant. In voting for a representative, which I consider next, citizens answer a question about who would advance the public good, and then the elected representative continues the citizens' thinking about that question, by organizing them within their own thinking, for use in a public deliberation. This use makes the citizens present in that deliberation, so that they are, derivatively, participants in it too.[70] Crucially, a representative can only make the thinking of citizens present to the national deliberation, and vice versa, by carrying out the deliberation publicly. Without publicity, the thinking

[68] McCain Counters Obama 'Arab' Question, https://www.youtube.com/watch?v=jrnRU3ocIH4.

[69] Thanks to Seana Shiffrin and Richard Re for asking about the scope of a representative's public activities.

[70] HANNA PITKIN, THE CONCEPT OF REPRESENTATION 63, 92 (1967) (discussing the idea of presence as an influence on views about representation).

of constituents dead-ends, as they have no way keeping up with the actions of their representatives and the interactive loop between representative and constituents closes.

In short, without publicity, citizens cannot hold representatives responsible for what the latter take responsibility for: public thinking. For this reason, representatives may not vote anonymously, and the reason is not that permitting it would do more harm than good (though that's true), but because doing so would be an abuse of anonymity. The purpose of representatives is to take responsibility for public thinking, which they do by thinking publicly and on behalf of the public. Consequently, voting anonymously would directly contravene their public purpose. Put differently, the duties of a legislator preclude the privilege of anonymity.

2. Citizens Have Non-Pragmatic Interests in Voting Anonymously

In this section, I argue that the secret ballot is favored for two distinct reasons, neither of which is pragmatic, in the sense of trying to forestall anticipated wrongdoing. I then consider the importance of these reasons. I argue, tentatively, that persons should be permitted to vote anonymously despite admittedly serious counter-considerations. The first reason I offer is one based in the interest of individual citizens, viz. citizens will often be, with reason, unready to take responsibility for their view of the public good, and yet they have an interest in providing it to the public. The second reason is a collective interest that citizens have in not holding each other responsible for how they vote, so as keep citizens from developing an excessively partisan identity.

The first reason for anonymous voting draws directly on the moral logic of anonymity; namely participation in public thinking. In the case of participation in the republic of letters, anonymity permits one to take advantage of the thinking of others before taking a position oneself. In interactions with the republic itself, the same logic is at work, though in a somewhat different form.

First, the structure, is different. When one reads or speaks anonymously to the republic of letters, one chooses the content one will speak or read. However, when citizens vote, the content is fixed. A citizen is asked a specific question, i.e., should this law be implemented or not, or: which of the following persons should hold office?

Second, the effect is different. Participation in the republic of letters does not exercise power over others in the same way as voting. The latter has legal effects, whereas the former, while by no means innocuous, has effects on what others consider or think about. Moreover, some ways of abusing anonymous speech are illegal, because the spoken content is independently wrongful. For instance, defamatory speech is illegal and so anonymous defamatory speech is too, and may be unmasked on that basis. Whereas, the use of anonymity to vote for damaging or self-serving policies is not, because persons are free to vote as they choose.

Third, the subject matter of voting is different and specialized, pertaining, as I will assume, to the public good.[71] That is, when one is answering a specific question in voting – the relevant metric in evaluating it is the public good. This assumption is controversial, but the controversy is innocuous here because it only makes my argumentative task harder. To the degree persons should be voting to maintain a private view of the good, then secret balloting will be even more justified.[72]

Given these points, we can see that open voting would require something fairly onerous of citizens; it would ask them to publicly answer a difficult question that they may not have settled to their satisfaction and that they wish to think more about. The difficulty of the question owes to the fact that it is about the public good, which is one that has been struggled with by countless thinkers.

---

[71] *Cf.* Annabelle Lever, *Mill and the Secret Ballot: Beyond Coercion and Corruption*, 19 UTILITAS 354, 369-70 (2007) (arguing for some caveats to the view that voting should be in the public interest).

[72] Then again, a counter-argument might be that if persons are supposed to vote their private view of the good, they may arrive at the answer more readily and so be expected to make it public, without invading their entitlement to keep thinking – as further thinking is unnecessary.

It's also intimidating because the lives of others are at stake. Moreover, voting requires applying one's theory of political morality and communal good (however developed) to the specific empirical circumstances one finds oneself in. One must know a fair bit about one's society and its conditions to understand what would move it closer to an ideal, flourishing, or fair one. These conditions are always changing. Furthermore, there may be strategic questions to consider, that arise from the way in which institutions aggregate votes, or how representatives will form coalitions or otherwise amass the needed political power. These present additional complications and are always changing as well. Altogether then, the question of who or what to vote for is one that persons may not easily answer, making them warranted in continuing to think about it.

By itself, the difficulty of who and what to vote for does not warrant anonymity for voting. One can simply decide not to answer. However, the option of answering may not be viable, either because exercising it will itself reveal one's thinking (refusing to answer some questions tends to reveal a certain answer to them) or because there is pressure to respond. Sometimes, there is pressure to answer that is not unreasonable. In law school, one may be asked a difficult legal question. All else equal, one may wish to think more about it, but the question comes with some pressure to answer. There are the social dynamics of being asked something in front of a group of one's peers, the need to evaluate students, and the need to elicit material for discussion and further contemplation as a group.

Public voting, it might seem, asks difficult questions to the electorate, but not with any real pressure, since person can choose not to vote or return an empty ballot. This result is arguably good. According to some, the whole point of public voting is to make persons pay more attention to how they are voting and so, to defer voting until they are confident enough to take responsibility for their judgment.

However, I think there is pressure to answer. Why? Because the vote of each person cannot wait. In a democracy, everyone's participation, on equal terms with others is a great good. If some feel comfortable voting their view of the public good while others wait to think more, then the former have, for the arbitrary reason that their thinking "finishes" first, an outsized presence and role in the formation of public life. Hence, the public loses out when persons don't vote, because the opportunity to educate and engage that person about civic affairs is deferred, and the opportunity to be educated by them is sacrificed as well.

These problems are bad enough as they are, but they are exacerbated by the fact that there are path dependencies in national life. Making some decisions at one point may put pressure, both rational and non-rational, to adopt other decisions later, so that not voting early may shut one out of fully participating in national life, given the later influence that the early votes have. For example, a decision at one time to pave park paths with gravel may provide some hindrance to later attempts to encourage bicycle use, since repaving the paths with smooth concrete may be prohibitively expensive.[73] A different kind of example is one in which voting by those confident enough to do so publicly in one election produce changes in society that are sufficiently significant such that those who would have been prepared to vote in the next election now feel that they must get to grips with the changes that have been wrought. The more cautious voters may be continuously playing catch up as the world moves around them. A third related example is that the state of the polity, its institutions and patterns of interactions, may tend to normalize or make salient certain kinds of conduct that dampen criticism or dissent. If persons wait to think more about how society ought to be, they may have to resist these kinds of pressures.

[73] *See* Eric A. Posner and Adrian Vermeule, *Legislative Entrenchment: A Reappraisal*, 111 YALE L. J. 1665, 1687 (2002) (giving this example).

It's true that those who are cautious about voting may realize that they must vote their best guess or be left out of national affairs, but this is not a good kind of pressure to apply. It frames the decision of whether to vote as one that is partly force by the thinking of others. This may breed resentment. A better solution is secret balloting, as this permits persons to vote their best guess and so to participate in politics right away and not to force them to choose between participation and continuing to think without taking responsibility for it. If this solution is warranted for some significant group of citizens, then there is a reason to make secret balloting the default for all citizens because if persons had to request to vote anonymously, then they would be trading one undesirable public stance for another. Requesting anonymity for oneself specifically would allow one to vote one's uncertain view of the public good, but one would then have to accept the suspicion of others that one has something to hide.

At this point, I will compactly review the argument and distinguish it from the prevalent view that secret balloting guards against voters being bribed or intimidated into changing their vote. In review, the argument is that elections, which have to be reasonably frequent, put a hard question to voters. They may, with reason, not be finished thinking about how to answer it. If voting is public, they must choose between participation or taking responsibility for a political view. If it is anonymous, they need not choose.

This view is a further application of my claim that the moral logic of anonymity (as namelessness, not non-identifiability) is to privilege participation in communal thinking. The right to speak anonymously, which is, doctrinally, given special protection in the context of political speech, is a way in which one can participate in the republic of letters even when one's thinking is uncertain, and voting anonymously is how one participates in the republic itself, when one's thinking about the public good is uncertain. Anonymous reading does double duty. It's not just how one receives

information from the republic of letters, but also from the republic. The former allows one to enrich one's thinking about a number of topics, whereas the latter allows the same, but where the topic is the distinct one of civic understanding or engagement.

This view is not the same as many going justifications for the secret ballot.[74] Many theorists think the reason for the secret ballot is to guard against wrongful attempts to influence voters. According to this view, persons should not have to vote publicly, because one effect of doing this will be to given an opening to those bad actors who wish to corrupt or intimidate voters. They will be able to see how persons vote, and so can, beforehand, credibly promise to pay for votes that they want cast or else issue threats against those they don't. So, the reasoning goes, to preclude these wrongful activities, secret balloting should be the norm.

By contrast, the rationale I am offering does not put the possibility of bad actors at the center of the case for the secret ballot. Rather, my view is that persons cannot be made to think with each other about politics because that subject matter is hard, and thus persons will very often not be finished thinking about it in time for the next election, and yet, their provisional political conclusions are needed. For this reason, the secret ballot is not a capitulation to a flawed and ugly world, but rather responds to a fundamental feature of thinking about the public good, as it finds expression in democratic politics. This is a better position than the alternative because it correctly explains the feeling that the importance of the secret ballot outstrips it prophylactic benefits, especially given that those benefits could likely be gained in other ways, e.g., aggressive investigation and enforcement of laws against voter intimidation and vote-buying.

---

[74] An exception is Annabelle Lever. She argues that anonymity is not an attempt to head off bad behavior, but, like me, a way of ensuring participation in government by citizens. Her argument is very different from mine though. She does not discuss anonymity's relationship to being responsible for thoughts and instead focuses on the risks of persons being shamed for how they vote if voting were open. Anonymity prevents this shaming on this view. Her argument is developed across two articles. Lever, *supra* note 71; Lever, *supra* note 50.

I have shown how a distinct justification for the secret ballot emerges from an application of its moral logic, that of being a privilege to participate in communal thought. As such a privilege, the secret ballot has significant similarities to anonymous speaking and reading. However, it is also dissimilar to these activities in that it has a collective dimension. I now shift my focus to an interest in anonymous voting that is had by the community at large rather than individual citizens. This interest is that of making salient the creation of "the people" as a collective subject of democratic politics.

The general source of anonymity's role in creating collectivities is that the import of its removal of an individual's identity can function differently depending on the background conditions. If an individual decides to make an anonymous statement, the statement has no owner. However, if an individual with several identities makes an anonymous statement, then the removal of the individual's identity may make room for a different one to take its place, or to enter salience. For my purposes, the kind of substitution that is most relevant is that of an institutional, collective identity rushing into the void left by a particular person's anonymity. In these cases, a refusal to possesses something as one's own is converted, via an institutional background, into a willingness to share responsibility with others.

A concrete example of this general phenomena is the *Economist's* editorial policy of not listing the authors of the articles it contains.[75] The purpose of this policy is to allow "many writers to speak with a collective voice."[76] A writer who produces an article does not take responsibility for her words, primarily, as herself, but as a member of a community. I say "primarily" to highlight an

[75] WHY ARE THE ECONOMIST'S WRITERS ANONYMOUS?, https://medium.economist.com/why-are-the-economists-writers-anonymous-8f573745631d (last visited Dec. 2, 2020).

[76] *Id.*

important wrinkle, which is that one can, for some articles, find out who wrote it without much trouble, because the *Economist's* website lists, at least, its major columnists.[77]

A second example is *per curiam* ("by the court") opinions. The practice is complicated and can have a range of forms and consequences.[78] However, its effect is at least sometimes to alter the structure of the resulting opinion. In many cases in the U.S., judges write for themselves, and the result is a composite of discrete viewpoints on the relevant legal questions, where the nature of the composite can be a matter of some complexity.[79] These opinions are of course "by the court," but only derivatively. They are so in virtue of being by the individual judges who make it up. However, in some situations, the disposition of the case by the court is not found by combining individual opinions, but rather, is presented fully formed and, so to speak, needing no assembly, in the name of the court itself. The institution itself speaks.

If we turn back to voting, we can see, with some differences, a similar paradigm at work: of individual identity effacement as tending to reinforce the salience and reality of a collective identity. When an election is held, the people speak, and the representatives become credentialed through that speech, in much the same way that a prosecutor gains the imprimatur of the people by proceeding on a grand jury indictment. In this way, the people serve as the "mind" of the polity. It embodies the shifting, tentative, and flexible view of the nation as it becomes expressed when the people speak in elections, and so casts their views in the form the government, with specific persons occupying its offices.

---

[77] ECONOMIST MEDIA DIRECTORY, https://mediadirectory.economist.com/ (last visited Dec. 2, 2020).

[78] *E.g.,* Stephen L. Wasby, *The Per Curiam Opinion: Its Nature and Functions*, 76 JUDICATURE 29 (1992); Richard Lowell Nygaard, *The Maligned Per Curiam: A Fresh Look at an Old Colleague*, 5 SCRIBES J. LEGAL WRITING 41, 46 (1995).

[79] Richard M. Re, *Beyond the Marks Rule*, 132 HARV. L. REV. 1942 (2019) (reviewing and analyzing the relevant doctrine).

Anonymity for individual voters preserves the flexibility of the people and their openness to new arguments, whereas when persons are publicly identified with their vote, the people becomes fragmented into particular people with particular views. This causes two kinds of inertia that get in the way of the people serving as a responsive addressee for democratic argument.

One is at the level of individuals becoming associated with a decision. If voters are publicly associated with their vote, they will develop a record, and records can create inertia in thinking. Persons may get into a rut, they may be pulled to vote again for the candidate that they voted for last time, or the policy they endorsed in the election before. Anonymity prevents the weight of past decisions from slowly calcifying the people into specific persons with known histories. In this way, the people has no history, no voting record, and no affiliation. It is recreated anew in each election to deliberate about the public good, *de novo*. Just as an anonymously published argument may have more force, owing to its depersonalized presentation, so maintaining the people as an abstract embodiment of the thinking of individual citizens, but not their particular characteristics or histories, helps preserve its flexibility and openness to argument.

By making this point about depersonalization, I am not going back on my earlier claim that the ideal for public deliberation is not anonymity but taking responsibility in one's own name. My point is rather that the dialogue between representatives and citizens works best when both poles are represented. Public deliberation is fulfilled by representatives taking responsibility for their votes, but it also well-initiated by preserving the people, as an experimental, thought-like body.

The second kind of inertia concerns the relationship between voters and their representatives. Representatives are supposed to govern for the people, which is to say, those who voted for them and those who did not. If persons were identified with their vote, interactions between constituents and representatives would be infused with a personal "with me or against me"

mentality. That is, politicians could know, when helping a particular constituent with an issue, or meeting to answer her questions, whether she voted for him or not, rather than reinforcing the voter's identity as a member of the people, and so deserving consideration on that basis.

Notice that my argument is not that anonymity can or should preclude persons from forming political identities. They may do so in a variety of ways, such as by being a member of a party and publicly supporting particular candidates as well as announcing their voting choices. Instead, my point is that secret balloting creates a starting point in which one is not affiliated with a position. Moreover, the secret ballot precludes there being an official record of a person's voting, even if he avows his electoral choices. This is justified, on the ground that doing so would constitute the state eroding the anonymity of others, even though this will have, in most elections, a negligible effect.

Having provided two considerations that support the secret ballot, I will now consider whether they should carry the day, all-things-considered. There is a serious argument that they should not, which is that persons must take responsibility for how they exercise political power. Even if my main claim is true, that the case for secret balloting is not based only in pragmatics, it may not be sufficient to justify it, all-things-considered.

I'm sympathetic to this objection. After all, it is one thing to read and speak anonymously in connection with the republic of letters, but political community is different. For example, though answering questions about the public good is difficult, citizens have help. A good democracy will foster an understanding of the role of civic institutions as well as thinking about collective problems faced by the community. As I argued before, representatives, agencies, and courts, not to mention think tanks and newspapers, provide examples of publicly-oriented thinking. Moreover, individual elections are preceded by campaigns, which provide further education about the relevant issues.

Last, it could be argued that a good democracy should require open voting, but cultivate a default tolerance of how others vote, rather than shield individual voters from taking responsibility for how they vote. Together, these reasons militate in favor of open voting, even if we accept the view of democracy according to which citizens answer questions about the public good.

There are reasons on the other side though too. One is that in voting persons exercise a symmetrical and equal political power to others, so that though one is not taking responsibility for how one votes, one's influence on an election overall is just as much as anyone else's. It's not generally true that persons must only take responsibility for actions they take in positions of power, but certainly, the exercise of power heightens the need for responsibility. A good leader must take responsibility for what happens in virtue of having prerogatives to make things go well as well as having an obligation to wield those prerogatives carefully. In an election among citizens, voting power should be equal. Given that, two citizens who disagree balance each other's vote. No doubt they should talk about their differences, but there many opportunities for them to do so, and in any case, a discussion mediated by other centers of thinking may be more productive. In the meantime, the exercise of power by these two citizens is reciprocal and frequently renegotiated through regular elections.

A second point about representatives that I have been emphasizing. Voting for a representative exercises power, but not for a final view of what the laws will look like. Hence, those who voted for the losing candidate may still attempt to persuade the elected representative to moderate her position on the relevant issues. Representatives represent all of their constituents, and so should stand in the way of the winning voters' attempt to take advantage of the losers.

Altogether, I think whether voting by citizens should be anonymous is a close case. It is best justified when voting for representatives, as anonymous voters can, in that case, serve as a flexible and experimental input to further deliberation.

3. Ballot Initiatives

The foregoing discussion helps answer the question of whether signing a ballot initiative should be something that citizens must do in their own name. It helps because it has illustrated, very broadly, one way that thinking "flows" in a democracy. Namely, a public question is posed to citizens who are thereby mobilized into a segment of "the people." When this question is sufficiently difficult, they are entitled to answer anonymously, and in so doing provide a proto-answer to that question, which is then taken up by a public official (prosecutor, legislative representative) as the initial step in her taking responsibility for fleshing out and enriching that answer, and to take, and take responsibility for, some corresponding public action, such as voting for the piece of legislation, or bringing charges.

Where do ballot initiatives fit in this cycle? They are hard to characterize because they involve ordinary persons, but also attempt to take legislative action. One complication is that some ballot initiatives do not receive enough signatures and so never appear on the ballot. To avoid this complication, and to focus on what I take to be the core, interesting case, I focus only on those ballot initiatives that garner enough signatures and otherwise succeed in being put on the ballot, such that a vote by the polity will determine whether the initiative becomes law. Even in these cases, a strong argument can be made that signatories to a ballot initiative should be entitled to anonymity, on the grounds that they are only suggested public actions, that are purposefully submitted to the public at large for approval. Due to their willingness to seek the public's approval for what they have

only suggested, it may seem that ballot initiative signatories should be allowed to be anonymous and so avoid being held responsible for their stance on the issue.

Despite this argument, I think that signatories to a ballot initiative undertake a public act, and so must do it in their own name. Not only is disclosure of the signatories permissible, but there is some reason to disclose them. This reason, to be unpacked, is that a ballot initiative is not a query of the people, for which an institution or official stands ready to take responsibility for the answer in further thinking. Put another way, ballot signatories wish to put an idea on the collective agenda that has, except for the signatories, no spokesperson. Hence, the signatories become the spokespersons.

The first point is that ballot signatories are not responding to a question that they might otherwise have wished to keep thinking about. Rather, they see the initiative as a good idea, or, at least, they wish others to consider it *now*.[80] They thus set the timing for the deliberation of others as well as the form it takes, i.e., whether the initiative will be enacted will turn on the deliberations of the members of the polity themselves and not their representatives. This shift in timing and form is not trivial, as the passage of the initiative may have follow on effects, such as giving legislators reason to take up different business than they otherwise would, to respond to the law that was passed, such as to limits its bad effects, or to make it consistent with other regimes or systems of regulation.

Second, anonymity, if it were to be granted to ballot initiative signatories, would not be securing a general participatory right for persons to equally consider a question, at their own pace, but would protect those who share a particular perspective and have organized on that basis. The supporters of the initiative have conducted their own, selective and informal polling of the public,

[80] *E.g.,* Gertz v. Robert Welch, Inc., 418 U.S. 323, 352 (1974) (arguing that petitioner Elmer Gertz was not a public person because he did not "engage the public's attention in an attempt to influence its outcome.").

where only those who support the initiative are recorded and those who did not sign are irrelevant, even if their number is comparatively large. This makes a ballot initiative a one-sided mobilization of the public, where supporters are gathered and opponents are filtered out.

This one-sidedness does not make ballot initiatives suspect or wrongful. Most of politics consist in organizing the like-minded, but the point is that an attempt to organize those who are like-minded to change the power of other citizens should be done in the name of those who wish to band together. They are not a manifestation of the people, but rather, are some particular people who found others who agree with them, and they should be identified as such.

Third, if the ballot initiative is successful, then the result is a new option to vote for, but the option does not come with any person who takes responsibility for it, in the sense of advocating it and being a target for questions and challenges regarding it. The electorate is thus on its own in trying to gauge its desirability. And not only that, those who signed the initiative have already started thinking about it. They are familiar with it and have been, in the ordinary case, been exposed to the benefits of the proposal by those who favor it. This combination – the lack of a spokesperson for the law and the comparative expertise of those who have endorsed its proposal, as compared to the ordinary member of the electorate – militates in favor of treating the signatories as stepping into the shoes of such a spokesperson. Persons may use the identities of those who signed the ballot initiative as a clue by which to assess the proposal. If the signatories are from a particular group or industry, this may provide information about how to go about researching the proposal itself.

A summary of the argument is that signatories to a ballot initiative are a subset of the public acting *sua sponte*, not responding to a question put to the public, but rather wanting to push their own agenda about what the public business should be. This is not a bad thing. Democracy is well-served by active citizens, but their activity as a particular subset of citizens is what requires publicity.

It is activism borne of a particular viewpoint and desire to move it forward that requires doing it in one's own name.

Let me note one limitation of the argument. I am not claiming that ballot initiative signatories are never entitled to anonymity. Rather, I agree roughly with the concurrence in *Doe v. Reed* in thinking that ballot initiatives are, by their nature, public, but that they may earn anonymity when a cognizable risk of reprisals or harassment can be shown.[81] The point is that this result is arrived at by an argument that has a different premise than the claim that all acts with legislative effect must be public, or that the exercise of political power must always be public. This premise is that only those acts for which one must take responsibility must be public. Since a ballot initiative signatory must, I argued, take responsibility for signing, her signature must be public, but this is compatible with there being no need to take responsibility for other acts that exercise political power.

IV. Conclusion

My goal has been to identify and replace a pervasive assumption in legal and philosophical thinking.

The assumption is that anonymity is just a condition that persons can be in, and that whether being in that condition is good or bad depends solely on the situation and what one wants to accomplish. This is the assumption that anonymity is not morally distinct. If one accepts it, then individual uses of anonymity will have to be analyzed case-by-case. With regard to each case, the question will be "what do we want to happen in this situation?" followed by an inquiry into whether

[81] John Doe No. 1 v. Reed, 561 U.S. 186, 215 (2010) (case specific relief is available "in the rare circumstance in which disclosure poses a reasonable probability of serious and widespread harassment that the State is unwilling or unable to control.") (Sotomayor, J., Stevens, J., Ginsburg, J., concurring).

anonymity will further that purpose. Using anonymity rather than some other technique, such as anti-retaliation measures, will be a purely technological one, turning only which can produce the best results, as one judges them to be.

I argued that this assumption should be replaced. Anonymity is not a morally indifferent state whose production should be judged solely by the results of producing it, but rather a condition that has, by its nature, moral implications. More specifically, anonymity permits persons to interact with ideas without being held responsible for how they do so. Thus, anonymity is an exception or a privilege from the normal order of things, in which persons may be held responsible for whatever they are responsible for. Such a privilege requires justification, and the justification is not wide open, consisting of expedience or of "goodness" however denominated, but is instead one that ensures the realization of the good *of anonymity*, as a distinct moral practice.

The good of anonymity, I have argued, is that it permits a form of thinking by which persons can retain the privacy of their thoughts without having to withhold them from others. With anonymity one, retains privacy in one's ideas, not by holding the ideas back, but by holding one's self back, so that others do not have the means of identifying the one who holds them. The justification for anonymity is to privilege participation in community intellectual life.

If we accept this replacement account of what anonymity is, then we will have to stop thinking about various anonymity practices as simply useful or not. Instead, we will have to ask, of an instance of anonymity, about whether someone should be held responsible for what they do anonymously, and this is a different question. It requires thinking about the allocation of responsibilities to persons – what should persons have to own and what may they disown? One answer is that in some circumstances, persons must disown their intellectual activities, as these are a part of getting situated for those activities which they must perform in full view of others.

Chapter 3: Thought Crimes

I. Introduction

In the previous two chapters I focused on the pitfalls of prematurely holding someone responsible for what she thinks, with some application to democratic practices. In this chapter and the following one, I look at legal doctrines in the criminal law that shield persons from being held responsible for what they think by punishing their thoughts (this chapter) or by legally requiring that they reveal their thoughts (the next chapter).

Imagine that a man sits in a restaurant and writes out, on a paper napkin, a personal resolution of sorts. In it, he enumerates his reasons for, and his consequent harboring of, an intention to attack a particular person with brass knuckles. He leaves without taking the napkin with him, and the server cleaning his table picks it up, reads it, and gives it to the police. There is thus evidence that the author has at least two thoughts with criminal content: he *believes* that he should, all things considered, use brass knuckles against another, and he *intends* to do so.[1]

Nonetheless, he cannot be convicted of any crime, whether for his belief or his intention, because of the "centuries old maxim" that "no one is punishable solely for his thoughts."[2] This venerable doctrine, the *thought crime doctrine*, is widely accepted, but only partially understood.[3]

[1] *See* Nicole Hong, *Ex-C.I.A. Analyst Faces Trial in Biggest Leak of Agency's History,* N.Y. TIMES, February 5, 2020, at A25 (former CIA employee's diary used as evidence of his intention, but used only to establish the mental element of a crime that also includes an act element); Gerald Dworkin & David Blumenfeld, *Punishment for Intentions*, 75 MIND 396, 400 (1966) (imagining the discovery of a diary that reveals a person's intention).

[2] United States v. $11,500.00 in United States Currency, 869 F.3d 1062, 1075 (9th Cir. 2017) (collecting sources and discussing the thought crime doctrine at length, with emphasis on intentions).

[3] ALAN BRUDNER, PUNISHMENT AND FREEDOM: A LIBERAL THEORY OF PENAL JUSTICE 108 (2009) (noting wide agreement that thoughts may not be punished).

As applied to beliefs, the doctrine has been well discussed and plausibly defended on First Amendment grounds.[4] According to these arguments, the freedom to believe, free from punishment, is necessary to realize a number of individual and political values. Arguments of this kind are often accompanied by the observation that beliefs do not aim at changing the world, but only at representing it. Consequently, beliefs are, in a limited but important sense, unthreatening to public order.[5] People may believe what they will, and yet, respect the law.

But the thought crime doctrine also includes intentions – understood here as commitments to act, either now or in the future – and their inclusion is both much more puzzling and much less discussed.[6] It is more puzzling because intentions need not have either of the features that make beliefs a good candidate for protection under the First Amendment. Intentions to commit crimes, for example, have no role in a robust free speech culture and they motivate criminal acts. Such

---

[4] *See, e.g.*, Nat'l Inst. of Family & Life Advocates v. Becerra, 138 S. Ct. 2361, 2379 (2018) (Kennedy, J., concurring) ("Freedom of speech secures freedom of thought and belief."); Abood v. Detroit Bd. of Ed., 431 U.S. 209, 234-6 (1977) (giving examples of the claim that "…freedom of belief is no incidental or secondary aspect of the First Amendment's protections…"); Stanley v. Georgia, 394 U.S. 557, 565 (1969) (discussing the right to receive various kinds of ideas, but also discussing the lack of a government interest in preventing "antisocial" thoughts); Speiser v. Randall, 357 U.S. 513 (1958); W. Virginia State Bd. of Educ. v. Barnette, 319 U.S. 624, 641 (1943) (discussing freedom of belief as necessary for consent to be governed); Cantwell v. State of Connecticut, 310 U.S. 296, 303-311 (1940) (discussing belief in a religion); Whitney v. California, 274 U.S. 357, 372-8 (1927) (Brandeis, J., concurring) (discussing belief in an ideology that favors lawbreaking).

[5] First Amendment precedent draws this distinction in the course of protecting beliefs. *See, e.g.*, Brandenburg v. Ohio, 395 U.S. 444, 448 (1969) ("the mere abstract teaching of the moral propriety or even moral necessity for a resort to force and violence, is not the same as preparing a group for violent action and steeling it to such action.") (quoting Noto v. United States, 367 U.S. 290, 297-8 (1961)); Stanley v. Georgia, 394 U.S. 557, 565 (1969) (holding that persons may "read or observe" what they please).

[6] *See* Douglas Husak, *Does Criminal Liability Require an Act?, in* PHILOSOPHY OF CRIMINAL LAW: SELECTED ESSAYS 51 (2010) ("When there is overwhelming evidence that a defendant firmly intends to commit a crime, why should punishment be undeserved simply because the act requirement is unsatisfied? I am unaware of a principled answer to this question.").

Only a small number of Supreme Court cases deal with the thought crime doctrine as applied to intentions, and those that do offer relatively sparse analysis. *See, e.g.*, United States v. Resendiz-Ponce, 549 U.S. 102, 107 (2007) ("Not only does the word "attempt" as used in common parlance connote action rather than mere intent, but more importantly, as used in the law for centuries, it encompasses both the overt act and intent elements."); United States v. Shabani, 513 U.S. 10, 16 (1994) ("…the law does not punish criminal thoughts…"); Robinson v. California, 370 U.S. 660, 678 (1962) (Harlan, J., concurring) (holding that a person may be punished for a "bare desire" to commit a crime, in a context where such a desire would arguably become an intention).

See also, Young v. State, 303 Md. 298, 309, (1985) ("it is not enough that a person merely have intended and prepared to commit a crime."); Michael L. Rich, *Limits on the Perfect Preventive State*, 46 CONN. L. REV. 883, 910 n.181 (2014) (collecting secondary authorities regarding the thought crime doctrine).

intentions thus challenge public order to a greater extent than beliefs without any compensating speech benefits.

Furthermore, intentions like those of the resolution author seem relevantly similar to many actions that are, apparently, justifiably criminalized. For instance, actions that correlate with and enable future wrongdoing may be punished as preparatory crimes, such as the possession of burglary tools or brass knuckles.[7] Actions that advance or further a wrong are punishable as attempts, such as putting on brass knuckles in order to attack with them.[8] And actions that aim to get another to commit a crime are punishable as solicitations.[9] But the resolution author's intention is not unlike these actions. It correlates with and enables an attack with brass knuckles. It also furthers such an attack. Moreover, if it is punishable to try to convince, suggest, or request that someone else use brass knuckles to attack a third party, then why is it permissible for the resolution author to harbor such an intention himself?

These puzzles set the stage for the main question addressed by this chapter: if criminal intentions (such as the resolution author's) have little to do with the exchange of ideas, and if they have features that support their criminalization, then why not relax the thought crime doctrine to permit their punishment as a kind of a super-inchoate crime? Is there any *principled* reason, i.e., one not grounded in doctrinal preference or pragmatism, for why the thought crime doctrine should

[7] *See, e.g.*, Commonwealth v. Hardick, 380 A.2d 1235 (Pa. 1977) (burglary tools); Commonwealth v. Jones, No. 3345OF2001, 2002 WL 34401140 (Pa. Ct. Comm. Pl. Oct. 08, 2002) (brass knuckles). *See also,* Husak, *supra* note 6 (using brass knuckles as an example).

[8] *See* Wayne R. LaFave, 2 Subst. Crim. L. §11.2 (3d ed. 2017) (summarizing attempt doctrine); Mitchell Berman, *Attempts, in Language and in Law*, 6 JERUSALEM REV. LEG. STUD. 1 (2012). *See also* United States v. Soto-Barraza, No. 15-10586, 2020 WL 253560, at *2 (9th Cir. Jan. 17, 2020) (discussing and summarizing attempt liability principles).

[9] *See* Wayne R. LaFave, 2 Subst. Crim. L. §11.1 (3d ed. 2017) (summarizing solicitation doctrine). *See also* United States v. Williams, 553 U.S. 285, 297-301 (2008) (applying solicitation principles to child pornography law); United States v. White, 610 F.3d 956, 960 (7th Cir. 2010) (discussing solicitation principles).

embrace intentions in addition to beliefs? If there is, it seems that it must stem from nothing but the fact that intentions are *thoughts*. But what's the significance of *that fact*?

Legal scholars have wrestled with this question, but they have not decisively answered it.[10] For example, R.A. Duff argues that intentions are ineligible for punishment because they are a kind of incomplete attempt (attempts that do not do everything necessary to complete the base crime), and incomplete attempts are generally ineligible for punishment. However, the latter claim is implausibly strong. Moreover, Duff's argument provides no distinct reason against punishing thought and begs the question by assuming that only complete wrongs are eligible for punishment.

In a recent article, Gabriel Mendlow argues, as a matter of principle and current law, that conduct may be punished only on the condition that it may be prevented or interrupted by force.[11] Thus, thoughts can't be punished because that condition is not satisfied – the government may not, for example, extinguish a thought with a mind scrambler or forced neuro-surgery. However, the relied upon link between punishment and disruption is controversial, and, if established, only guarantees *that* thoughts may not be punished, but still does not give the justification for this conclusion.

I will offer a different defense of the thought crime doctrine, which is that it protects thinking. I have already defended the major premise of this argument, which is that thinking is

[10] *See* Gabriel Mendlow, *Why is it Wrong to Punish Thought?*, 127 YALE L.J. 2342 (2018); R.A. Duff, *Risks, Culpability, and Criminal Liability, in* SEEKING SECURITY: PRE-EMPTING THE COMMISSION OF CRIMINAL HARMS 121 (G.R. Sullivan and Ian Dennis eds., 2012); R.A. DUFF, ANSWERING FOR CRIME, 102-5 (2009); R.A. DUFF, CRIMINAL ATTEMPTS, 188-190; 386-393 (1996).

For additional scholarship about thought crimes, *see* Federico Picinali, *A Retributive Justification for Not Punishing Bare Intentions or: On the Moral Relevance of the 'Now-Belief'*, 32 LAW AND PHIL. 385 (2013); Husak, *supra* note 6; Clay Calvert, *Freedom of Thought, Offensive Fantasies and the Fundamental Human Right To Hold Deviant Ideas: Why the Seventh Circuit Got It Wrong in Doe v. City of Lafayette, Indiana*, 3 PIERCE L. REV. 125 (2005); Dworkin & Blumenfeld, *supra* note 1.

[11] *See* Mendlow, *supra* note 10, at 2346 (introducing this claim).

valuable for persons, so I focus on the minor premise, which is that beliefs and intentions are constitutive of thinking, and hence to criminalize them is to objectionably criminalize thinking itself.

The chapter proceeds as follows. In section I, I look at the theories of Duff and Mendlow. I argue that neither one offers a satisfying account of the thought crime doctrine. Moreover, Mendlow's characterization of current law is called into question by considering the prior restraint doctrine, injunctive relief, and criminal omissions. In section II, I turn to developing my own argument in two premises: that persons have an interest in thinking about what is true and that, given the structure of thinking, punishment for beliefs violates that interest. In section III, I show how a similar argument can be given for intentions and then defend it against some possible objections. I conclude by showing that the view I offer avoids the objections raised against Duff and Mendlow.

II. Contemporary Thought Crime Scholarship: R.A. Duff and Gabriel Mendlow

A. Duff

Across several publications, R.A. Duff develops a theory of criminal attempts, according to which early stages of criminal action, including but not limited to merely harboring an intention to commit a crime, are not punishable.[12]

The motivating idea for his view is that persons should not be punished merely on the basis that they will, in fact, go on to commit a crime.[13] Science fictional scenarios attest to the plausibility of this intuition, by imagining extraordinary means by which the government could look into the

[12] *See supra* note 10.

[13] DUFF, CRIMINAL ATTEMPTS, *supra* note 10, at 190 (persons should be subject to criminal sanctions "only if or when they have definitively refused to recognize or to act in accordance with those reasons: for if they were to be made liable to sanctions because of what they *would* do or *might* do, rather than because of what they have actually done, they would no longer be treated as rational agents who can, and should therefore be allowed to, determine their own conduct.").

future and accurately predict criminal activity before the "defendant" has even conceived of the crime.[14] Such "future wrongs" it seems, may not be the basis of punishment.[15]

The reason that future wrongs may not be the basis of punishment, Duff thinks, is that persons are responsible for their actions: they are, in principle, capable of seeing reasons and so acting rightly, even when they are, in practice, so set on wrongdoing that they will go through with it.[16] Thus, if persons are to be respected as being, in principle, capable of right action, then the government must withhold punishment until a person has *actually* acted badly and thereby made themselves, in principle, unable to act rightly.[17] It's not enough to predict that the person will act badly at some point in the future. The conclusion that Duff draws is that persons cannot be punished for intending to commit a wrongful act, because the intention is not sufficient evidence that they will commit it.

I agree that persons should not be punished *merely* for what they will or will likely do, but only for they have done. However, this fact – that punishment must be for what a person has actually done – does not automatically condemn thought crimes, because arguably, a person who intends to commit a crime *has* done something wrong – he formed an intention to commit a crime.[18] Why can't he be punished for that?

Let me put this challenge a second way. Remember, one motivation for thinking that persons may not be punished for future wrongs is that otherwise, someone who had not yet even

---

[14] *See, e.g.*, PHILIP K. DICK, MINORITY REPORT AND OTHER CLASSIC STORIES (Citadel 2016) (1956) (in which "pre-cogs" can see the future and so predict a person's crime before he even thinks to perpetrate it).

[15] There is a vigorous legal debate about punishment or at least confinement for future crime. *See, e.g.*, Richard Lippke, *No Easy Way Out*, 27 L. & PHIL. 383 (2008).

[16] DUFF, CRIMINAL ATTEMPTS, *supra* note 10, at 388.

[17] *Id.* at 190 (the government may punish persons "only if or when they have definitively refused to recognize or to act in accordance with those reasons: for if they were to be made liable to sanctions because of what they *would* do or *might* do, rather than because of what they have actually done, they would no longer be treated as rational agents who can, and should therefore be allowed to, determine their own conduct.").

[18] Dworkin & Blumenfeld, *supra* note 1, at 400 (making a similar point, i.e., that it begs the question to argue that thoughts cannot be criminalized because the person has not yet committed a non-thought crime).

considered committing a particular crime could be punished, so long as it were true that he would, predictably, commit it. And that seems like punishing a person while he is still innocent. But someone who intends to commit a crime is not innocent in this way. Such a person *has decided* to commit a crime. Thus, punishment for this person would not be *merely* for what he will do, but also for what he has already done, which is to decide in favor of wrongful action.

Duff's response to this problem – that his initial principle about punishment does not rule out thought crimes – is to supplement his principle. Instead of claiming that persons may only be punished for the wrongs that they have committed, Duff's supplemented principle claims that persons may only be punished for the *substantive* wrongs that they have committed, such as arson or attacking another with brass knuckles (and not for "non-substantive" wrongs that only aim at a wrong, like attempted arson, or intending to attack another with brass knuckles).[19] His defense of this supplementation is that someone who intends wrongdoing could still, in principle, be persuaded to give up his criminal enterprise before the substantive wrong is completed, and that treating him as free requires giving him this chance.[20]

Giving him this chance seems overly indulgent though. The fact that he intends to do wrong is some evidence that he will not be persuaded by the relevant reasons. Moreover, an intention to do wrong is a commitment to act, which may, in some cases, entail resistance to or avoidance of attempts at persuasion. Besides, waiting to see what happens imposes costs on others who may feel fear, or need to prepare to repel the wrong or minimize its damage. It is not persuasive to require these burdens for the flat justification that we need to treat the intending party as someone who can

[19] *See* DUFF, CRIMINAL ATTEMPTS, *supra* note 10, at 388 ("Suppose we know that someone intends to commit, or is preparing to commit or taking initial steps towards committing, a substantive crime. If we have the moral standing to intervene…we can properly do so by trying to dissuade him from continuing his criminal enterprise… or perhaps by reminding him of the prudential reasons for desistance provided by the threat of punishment for the substantive crime.").

[20] *Id.*

possibly give up his enterprise. He *can* give up this enterprise, but what's so important about treating him as if he can, and can't we treat him that way by just believing that he could give it up, but nonetheless suspect he will not and so punish him?[21]

An additional problem with Duff's supplemented principle is that it does not offer any *distinct* reason not to punish thought, and thus does not justify only the thought crime doctrine, but rather entails a broader "non-substantive wrong" doctrine. But that broad doctrine is implausible, as it requires doing away with a worryingly large chunk of attempt liability.[22] For example, Duff's supplemented principle seems to have the counterintuitive consequence that someone who is about to attack another with brass knuckles cannot be punished because he has not yet committed a substantive crime. After all, he could, in principle, abandon the attack at the last second.[23]

Duff tries to back away from this counterintuitive consequence by arguing that his position would not entail the complete abolition of attempt liability. It would not, according to Duff, because someone "in the process of committing" a crime, i.e., about to attack, has "crossed the Rubicon."[24]

Consequently, punishing someone for being *in the process of* committing a crime (where being in the process of committing a crime involves being committed to it) would not "fail to treat him as a responsible agent."[25] But these comments are an *ad hoc* response. As Duff admits, someone about to attack is still capable, in principle, of not attacking, and so seems to be immune from punishment

---

[21] Thanks to Richard Re for pushing me to make these arguments more explicit.

[22] DUFF, CRIMINAL ATTEMPTS, *supra* note 10, at 389 (acknowledging that his argument appears to have this consequence before trying to show that it does not). Importantly, Duff's doctrine would preserve attempt liability for someone who does all that they can do to bring about the result and nonetheless fails, e.g., shooting but missing. Still though, many attempts are intuitively wrong well before they reach the last act before the completion of the crime, which is my point in the text.

[23] *Id.* at 190 ("We may be certain that someone *will* break the law, that he will not be dissuaded from doing so by the reasons which the law offers him: but we must still treat him as someone who could be rationally persuaded to obey, even at the last moment."). Again, it should not be overlooked that Duff's principle would permit punishing persons who have done all they can to bring about an illegal result. So, someone who swung with brass knuckles but missed could be punished under his theory.

[24] *Id.* at 389-90.

[25] *Id.* at 390.

according to his principle.[26] To address this problem, he shifts his theory a bit and points out that someone in the process of committing a crime is *committed* to it, even if they could abandon it before doing all that was necessary to complete it. But if being committed to a substantive wrong can warrant punishment, then we are back at square one, because intentions to commit a substantive wrong are commitments to do so.[27] So, they should be punishable too.

To sum up: Duff's ultimate claim is that persons cannot be punished as long as they have not yet committed a substantive wrong. But no reason is given for limiting the principle to substantive wrongs. Moreover, the principle is implausibly broader than the thought crime doctrine, in that it prohibits the punishment of attempts.

B. Mendlow

In a recent article, Gabriel Mendlow offers a deductive argument for the conclusion that thoughts may not be punished.[28] The driving idea behind the argument is that criminal punishment and crime prevention go together, such that if an activity may be punished for being wrong, then it may also be disrupted for being wrong.[29] Mendlow calls this idea the *enforceability constraint*, and he thinks it is both a sound moral principle that all legal systems should follow (like "punishment

[26] *Id.*

[27] This is an intuitive point, as one can desire to do something, but not do it, because the costs are too high. An intention however is a decision to act and thus commits one to acting when the time comes. The idea that intentions are commitments to action is also a standard philosophical point. *See, e.g.*, Pamela Hieronymi, *Controlling Attitudes*, 87 PAC. PHILOSOPHICAL Q. 45, 56 (2006).

[28] Mendlow, *supra* note 10.

[29] *Id.* at 2370 ("it's wrong for the state to punish offenses of a given type if it's always wrong in principle for the state to forcibly disrupt such offenses merely on the ground that they're censurable transgressions.")

should not be excessive"),[30] and an accurate statement of current U.S. law (i.e., that current doctrine adheres to the enforceability constraint, though obviously not under that name).[31]

The enforceability constraint is plausible and certainly holds true for a number of activities. Take, for example, selling loose cigarettes.[32] Persons can be punished for selling loose cigarettes and police officers may disrupt their sale by seizing them from a seller's hands. And it would be strange if they could not do so. After all, if selling loose cigarettes is sufficiently wrong to warrant punishment if it occurs, then it seems that police officers are warranted in keeping it from occurring in the first place, just for being wrong.[33]

If Mendlow is right that the enforceability constraint is a sound, general, legal principle, then it can anchor an argument against punishing thoughts. Remember, according to the enforceability constraint, if the government may punish an activity for being wrong, it may also disrupt it for being wrong. So, if the government may punish (a person for having) a wrongful thought, then it must be true that it may disrupt that thought just because it is wrongful. But, Mendlow argues, the government is absolutely barred from disrupting thoughts, such as by administering thought-suppressing drugs or by employing a science-fictional thought scrambler.[34] Therefore, the government may not punish thoughts either.

---

[30] *Id.* at 2371-73.

[31] *Id.* at 2367, 2373 ("Justifying the Enforceability Constraint more fully is beyond the scope of this Essay. My present goal is more modest. It's to show how abnormal it would be to treat any type of transgression as an exception to the Enforceability Constraint….no recognized limit on the state's enforcement power does more than restrict when, how, or pursuant to what procedures given instances of an offense may be forcibly disrupted.").

[32] *Id.*

[33] The "just for being wrong" part is needed to rule out other grounds on which a person's thoughts could be extinguished or manipulated by the government, such as to prepare a defendant for trial. The fact that such mental disruptions are allowed does not warrant punishment for thought, because they are not undertaken on the grounds that the target thoughts are wrongful. Rather, the target thoughts are compulsive or otherwise deleterious to the person's mental competency. *Id.* at 2376-9.

[34] *Id.* at 2368, 2376-2383.

Another way to put the argument is as follows. The government may punish the sale of loose cigarettes on the condition that it may seize them from one's hands. And it may seize them, so punishment is permissible. Similarly, the government may punish thoughts on the condition that it may "seize" them, i.e., erase them from one's mind. But it may not seize them, so punishment for thoughts is impermissible too.

Despite its plausibility in a number of cases, the enforceability constraint is flawed. I will argue, in the following two subsections, that it is false as a statement of current law, and that it is doubtful as a justification for the thought crime doctrine.

1. The Enforceability Constraint is False as a Description of Current Law

Contrary to the enforceability constraint, current legal doctrine recognizes conduct that may be punished, and yet, not forcibly disrupted. Two examples are (pure) criminal omissions and the prior restraint doctrine.

Some crimes can be committed by omission or by commission. For instance, one can criminally trespass by going onto property or refusing to get off of it, because these are both ways that one's body can be in the wrong place at the wrong time. For crimes like these that may, but need not be, accomplished through omissions, disruption in Mendlow's sense is possible because the state of affairs that results from omitting to act can be prevented through force: a police officer can disrupt a criminal trespasser who omits to leave another's property by dragging him away.

However, some crimes are not a matter of some state of affairs obtaining (which the perpetrator caused or could prevent by acting), but necessarily a matter of failing to act. Such are what I call pure criminal omissions. An example is failing to file a tax return. This crime is, essentially, one of failing to do something; of not engaging one's will. Hence, it cannot be disrupted

by doing something to the perpetrator's body (or property or possessions), but only by goading the will itself. And current legal doctrine does not permit any such goading. For instance, police could not physically intimidate someone so that they file their taxes. Nor could they hypnotize them or trick them into filing. Even an injunction to require that they file a return would be unavailable, because generally speaking, equitable remedies like injunctions cannot be used to require persons to obey criminal laws.[35]

Failing to file a tax return is thus a duty that persons must undertake on pain of punishment, but not one they can be compelled to fulfill through disruptive force, because the only relevant kind of compulsion is the manipulation of a person's will. And the latter, *contra* the enforceability constraint, is illegal.

A second counterexample to the enforceability constraint, construed as a statement of current law, is the prior restraint doctrine. According to this doctrine, speech which may be later criminalized may not be disrupted ahead of time.[36] For instance, the government may not confiscate or destroy obscene materials in order to prevent their publication, even if they could be the basis for a criminal conviction after publication.[37] In these cases, punishment is permissible but not disruption.

---

[35] *See, e.g.*, United States v. Guest, 383 U.S. 745, 769 (1966) (Harlan, J., concurring) (citing the principle in passing but not applying it); United States v. Santee Sioux Tribe of Nebraska, 135 F.3d 558, 565 (8th Cir. 1998) (applying this principle); S.E.C. v. Carriba Air, Inc., 681 F.2d 1318, 1321 (11th Cir. 1982).

[36] *See* Near v. Minnesota, 283 U.S. 697, 713-14 (1931) (foundational prior restraint case in which an injunction against a newspaper was struck down); Martin Redish, *The Proper Role of the Prior Restraint Doctrine in First Amendment Theory*, 70 VA. L. REV. 53 (1984) ("Under the prior restraint doctrine, the government may not restrain a particular expression prior to its dissemination even though the same expression could be constitutionally subjected to punishment after dissemination.").

[37] *See* Quantity of Copies of Books v. State of Kansas, 378 U.S. 205, 210 (1964) (holding that seizing books without a judicial determination that they were obscene was a prohibited prior restraint); Channel 10, Inc. v. Gunnarson, 337 F. Supp. 634, 637 (D. Minn. 1972) (seizure of film was a prior restraint because it physically prevented plaintiffs from distributing it). *See also*, Gaetani v. Hadley, No. 14-30057-MGM, 2016 WL 593496, at *3 (D. Mass. Feb. 12, 2016) (noting that prior restraints of a purely physical nature are rare).

Mendlow anticipates this counterexample, and he meets it with two claims, one conceptual and the other legal.[38] Conceptually, he claims that injunctions are forms of disruption that are imposed solely for the wrongfulness of the enjoined conduct. He then adds the legal claim that speech injunctions are permitted so long they are supported by a finding that the relevant speech is not protected by the First Amendment. If he is right about both of these claims, then the prior restraint doctrine does not falsify the enforceability constraint after all. It does not forbid the disruption of punishable speech outright, but only requires that it take a certain form; namely, an injunction premised on a constitutional finding.[39]

Both components of this defense, however, are suspect. The legal claim is more ambitious than precedent will support. The case he relies on, *Balboa Island Village Inn v. Lemen*,[40] upheld a speech injunction for defamation, but not as punishment (only as a civil remedy), and moreover, only as an extraordinary remedy that was warranted by the defendant's willingness to repeat speech found to be defamatory. It expressly refused to hold that such speech could be preliminarily enjoined before it was uttered for the first time.[41] Indeed, such preliminary injunctions are generally treated as illegal.[42] Thus, if wrongful speech has not been uttered yet, it may not be enjoined and so not disrupted. Yet it may be punished (or at least subject to civil liability).

---

[38] Mendlow, *supra* note 10, at 2374-5.

[39] Mendlow, *supra* note 10, at 2375.

[40] 156 P.3d 339 (Cal. 2007).

[41] Balboa Island, 156 P.3d 339, at 344, 350, 351 (persons must be free to speak even if what they say will earn punishment after the fact, but making an exception for repeating such punishable speech).

[42] Mark A. Lemley & Eugene Volokh, *Freedom of Speech and Injunctions in Intellectual Property Cases,* 48 DUKE LAW JOURNAL 147, 149 (1998); Int'l Soc. for Krishna Consciousness of Berkeley, Inc. v. Kearnes, 454 F. Supp. 116, 122 (E.D. Cal. 1978) (Prior restraints are especially unjustified when the speech at issue is not yet been created, unlike obscenity in which the material at issue may already exist); Vrasic v. Leibel, 106 So. 3d 485, 486 (Fla. Dist. Ct. App. 2013) (denying a preliminary speech injunction because money damages were not shown to be inadequate).

A second source of trouble is Mendlow's conceptual claim that an injunction is a (i) disruption of the enjoined conduct (ii) for the reason that it (the conduct) is wrong.[43] No doubt injunctions are "forward looking"[44] in that they aim at conduct that has not taken place yet, but this is a far cry from showing that they have both of these two critical features.

Against (i), notice that injunctions themselves do not apply force to prevent an ongoing or future activity. They are not like seizing cigarettes from someone's hand. Instead, they are orders to act in a certain way. And while it is true that if the defendant does not obey the court's order, coercive force will follow, this fact does not mean that injunctions disrupt wrongful conduct in Mendlow's sense. After all, criminal laws work the same way. They announce a requirement on conduct and then impose penalties for its violation. Yet criminal laws do not disrupt.

Against (ii), it does not seem that injunctions (even preliminary injunctions) are disruptions of wrongful speech on the ground that the speech is wrong. It's true that such injunctions are properly *issued* for the reason that the enjoined speech is wrongful, but the coercive force that is applied for violating such an injunction is not for the fact that the violative speech was wrongful in itself, but rather because it constitutes disobedience to the court's order.[45] For example, if the defendant in *Balboa Island* were to violate the injunction against repeating the defamatory statements that she had made before, she would be punished with (criminal) contempt of court.[46] But, crucially,

---

[43] *See supra* note 33 (explaining why the enforceability constraint requires disruption to be engaged in for the reason that the disrupted conduct is wrong).

[44] *See* United States v. W. T. Grant Co., 345 U.S. 629, 633 (1953) (injunctions are forward-looking).

[45] An objector might interrupt at this point to stress that the punishment for the injunction might be for disobedience, and yet *also be* for the wrong itself, on the grounds that the wrong itself is the cause for the injunction being properly issued and so for its proper enforcement as well. Two possible responses are as follows. First, even preliminary injunctions are not justified simply because foreseen conduct is wrong, but, at least partially, to preserve the status quo for adjudication. *See* Roe v. Dep't of Def., 947 F.3d 207, 231 (4th Cir. 2020). Second, and similarly, the collateral bar rule suggests that an injunction is to be obeyed not because it is correct in addressing a wrong but because it is a kind of judgment that must be given authoritative force in order for justice to be administered in an orderly fashion. *See* In re Establishment Inspection of Hern Iron Works, Inc., 881 F.2d 722, 726 (9th Cir. 1989).

[46] Balboa Island, 156 P.3d at 353 (Baxter, J., concurring) (explaining that violation of the injunction at issue would lead to criminal contempt).

the basis of contempt would not be that she defamed the plaintiff (again), but rather that doing so would constitute a contemptuous attitude toward a court order telling her to stop. Consequently, her punishment would be for the wrong of disobedience, i.e., showing contempt, and not for defaming the plaintiff.[47] Criminal punishment, by contrast, is not imposed for disobeying a specific and explicit order to abstain from a wrong, but just for committing the wrong.

Given these weaknesses in Mendlow's response to the prior restraint objection, I conclude that it is, in addition to pure criminal omissions, a point at which current law diverges from the enforceability constraint.

2. The Enforceability Constraint Does Not Justify the Thought Crime Doctrine

Mendlow argues that enforceability constraint does not just describe current legal doctrine, but proscribes its correct form. The law should adhere to it because it is a sound moral generalization.[48]

I am unsure that it is sound, partly because it seems to me that the prior restraint doctrine is not arbitrary. It seems to have something going for it as a moral principle – that persons may sometimes be entitled to "lay what sentiments he pleases before the public" though doing so may incur punishment.[49] For instance, persons may be punished for not telling the truth under oath, but I doubt they can be given a truth serum to disrupt their perjury.

But let me stipulate that the enforceability constraint is a sound moral generalization in order to argue for a different point: that it does not constitute a justification for the thought crime

[47] *See* Margit Livingston, *Disobedience and Contempt*, 75 WASH. L. REV. 345, 351–2 (2000) (discussing reasons for various kinds of contempt penalties).

[48] Mendlow, *supra* note 10, at 2371-3.

[49] *See* 4 WILLIAM BLACKSTONE, COMMENTARIES *151–2.

doctrine. To set up this objection, consider this true moral generalization: it is impermissible to levy a $2 tax on some activity if it is impermissible to levy a $1 tax on it. Still, the truth of this generalization does not constitute a justification for the immunity that the activity enjoys from being taxed. Why is it impermissible to tax it even $1? The generalization just given assumes an answer to this question rather than giving it.

Similarly then, the claim that it is impermissible to burden thoughts with punishment if it is impermissible to burden them with disruption could be true, but it is not a justification for the thought crime doctrine, because it does not tell us why thinking enjoys immunity from being burdened. Even accepting this generalization, we still want to know: why is it impermissible to combat thoughts, even by disruption?

Mendlow does not focus on answering this question, but he addresses it briefly, suggesting that thoughts may not be burdened by disruption or punishment, because persons have the right to *mental integrity,* which is roughly, a right to retain whatever thoughts or mental states one has, free of brute psychological or physical interventions.[50] However, I doubt mental integrity is the base justification for the thought crime doctrine. My main reason for doubt is that it seems that there must be a (further) reason why persons are imbued with a right to have whatever thoughts they have.

To sharpen this argument, consider taxing thoughts. I submit that it would be wrong to tax thoughts based on their content (e.g., a tax for unpatriotic thoughts), and it is plausible that the reason for this is, at root, the same as the reason they cannot be disrupted or punished. Yet taxing something is not necessarily premised, as disruption and punishment are, on the illegitimacy of the

[50] Mendlow, *supra* note 10, at 2383.

thing being taxed.[51] A tax on thoughts would burden them, but not in a way that struck at a person's right to mental integrity, because a tax assumes the rightful possession or continuation of the thought, conditional on the payment of the tax. Hence, the right to mental integrity is likely based on a more fundamental reason that thoughts may not be burdened.

III. A Thinking-Based Argument Against Punishing Beliefs

In this section, I offer a justification for the thought crime doctrine in terms of the role that thoughts play in thinking: they are (a part of) one's thinking. Whereas Duff argues that thoughts are ineligible for punishment, a thinking-based view claims that their unrestricted formation is too valuable to the continuation of one's thinking to be punished, despite being eligible for it. Whereas Mendlow argues that the thought crime doctrine is justified by a right to mental *integrity*, a thinking-based view argues that it is justified by the value of mental *activity*: thinking.

I will introduce the thinking-based view through an argument focused on beliefs. It has two premises. The first is that persons have a preeminent interest in thinking about what is true. The second premise is that forming beliefs, even mistaken or wrongful ones, is an integral part of thinking about what is true. In the next section, I extend this argument to intentions by showing that they too are integral to thinking about what is true, though in a different way than beliefs.

[51] This point is a conceptual one that also enjoys legal acceptance. A tax does not entail that the taxed activity is prohibited or may be prohibited, but rather that it is conditionally, permissible, i.e., permissible on the condition that the relevant tax is paid. Supreme Court opinions agree. *See* Nat'l Fed'n of Indep. Bus. v. Sebelius, 567 U.S. 519, 573 (2012) ("the power to tax is not the power to destroy while this Court sits") (internal quotations omitted).

A. Premise One: Persons Have an Interest in Thinking

The major premise of my argument appears ubiquitously in First Amendment jurisprudence: that persons have an interest in thinking for themselves.[52]

I defended my reasons for embracing this premise in Chapter 1. To refresh that line of argument very briefly, I argued that the value of thinking should not be confined to its instrumental value in capturing the truth, just as the value of climbing a mountain should not be confined to reaching the top. Instead, just as climbing (as opposed to helicoptering) one's way to the top of a mountain is an activity that has its own distinct value, even when one doesn't make it, so the same for thinking. Thinking one's way to the truth has its own distinct value, even when one doesn't make it there.

Evidence for this claim is that thinking is valuable even when one has made or will make a mistake. Inquisitiveness is valuable by itself, and paternalism is forbidden, even though it paradigmatically involves "better" thinking by the paternalist. On the other hand, believing the truth without thinking one's way to it, by being programmed to have it is not worth much,[53] and attempting such a programming of another is seriously wrong. Similarly, it's a sensitive matter when one may rest a belief on someone else's say so. It's better to try to see the reasons oneself, though there may be practical limitations to doing this for every subject matter; there's just too much to think about.

---

[52] *See, e.g.*, SEANA VALENTINE SHIFFRIN, SPEECH MATTERS (2014); C. Edwin Baker, *Autonomy and Free Speech*, 27 CONST. COMMENT. 251, 259 (2011); David A.J. Richards, *Constitutional Legitimacy, the Principle of Free Speech, and the Politics of Identity*, 74 CHI.-KENT L. REV. 779 (1999); Charles Fried, *The New First Amendment Jurisprudence*, 59 CHICAGO L. REV. 225 (1992); David Strauss, *Persuasion, Autonomy, and Freedom of Expression*, 91 COLUM. L. REV. 334 (1991).

[53] Fried, *supra* note 52, at 232 ("no conviction forced upon us can really be ours at all").

Thinking is valuable because it is the performance of an obligation one has an interest in having and discharging, which is to try and keep one's own mental life in order.[54]

B. Premise Two: Punishing Beliefs Violates the Interest Persons Have in Thinking

I will now argue that punishing beliefs – I will focus on the punishment of *false* beliefs[55] – is seriously objectionable because doing will necessarily violate a person's interest in thinking. The reason, roughly speaking, is that carrying on the activity of thinking partly consists in making judgments about what is true *and* believing them. Judging and believing are two ways of thinking something is true, the former at a single time, and the latter over time. So, in the case where someone has judged that something is true and then believes it on that basis, punishing their belief amounts to punishing them for thinking (just as much as punishing them for making a judgment would). I will fill out these remarks with a general discussion of how judging and believing make up a paired unit of thinking, and then turn to a concrete example. One word of forewarning: in what follows, I often use thinking to encompass to making a judgment and the preservation of that judgment by believing it.

Judgments and beliefs form a rational couplet or a two-step. If one judges that something is true, then one should, rationally speaking, believe that it is true. If one judges that global warming is caused by humans, then one should believe that it is. The first takes place at a moment and the second is the initiation of an ongoing activity – if one believes something, one will keep believing it

[54] *E.g.*, 12 ANGRY MEN (1957) (Juror number 8, Henry Fonda's character, represents the value of thinking by insisting the other jurors at least think earnestly about their votes, even while admitting he may not know the truth.).

[55] I focus here because it's easier, I take it, to justify punishing false beliefs than punishing true ones. Hence, I make my argumentative burden harder by focusing on the former case.

unless sufficient doubt is cast on it.[56] It's helpful to think of the second, of believing the content of a judgment, as its psychological preservation or memorialization, the way that writing down the content of a judgment would constitute its written memorialization.[57] The key claim about this couplet that I wish to defend is that both parts of it, both judging and believing, are related ways of carrying out the activity of thinking, with believing being an enrichment of judgment.

Take judging first. It will, I think, be readily accepted that judging is a part of thinking. Indeed, it's a fundamental unit of thinking, because there is no way for one to know the truth unless one attempts to reach a bit of it, and a judgment is such an attempt. It's a shot at being correct about one bit of the truth. It's just a shot though. Judgments are fallible, and since thinking runs on judgments, it is too. As we know, mistaken thinking is common. Nevertheless, a judgment is, as a bit of thinking, not defective just because it is wrong, just as an attempted free throw is not defective, as a practice shot, just because it misses. In other words, a mistaken judgment is just as much a bit of thinking as a correct one.

Moreover, thinking is not reducible to one's judgments. One's thinking is not only as good as one's worst judgment. The reason is that thinking does not just take judgments as fixed, but seeks to strengthen them by subjecting them to pressure, putting them in conversation, learning from past misjudgments, and accounting for the perspectives of others.[58] Put another way, thinking uses judgments, but it uses them advisedly, with its eyes open about their weakness, in the service of transforming and structuring them into a more accurate and defensible whole. We might call this transformation theorizing.

---

[56] *See* Matthew Boyle, *Two Kinds of Self-Knowledge*, 78 PHIL. AND PHENOMENOLOGICAL RES. 133 (2009) (discussing believing as an ongoing process); Pamela Hieronymi, *Believing at Will*, 35 CANADIAN J. OF PHIL. 149, 174 (2009).

[57] *See* Tyler Burge, *Content Preservation*, 102 PHIL. REV. 457 (1993) (discussing memory as carrying forward a past judgment in time, and making it present now)

[58] Practicing a basketball shot or a tennis stroke is way of reflecting on and incorporating past shots and strokes to build up greater skill.

This last, transformative, aspect of thinking is where believing comes in. Beliefs make such transformations possible. Making a judgment is a one-time thing. If persons only made judgments, thinking would be difficult and might be impossible, consisting as it would of a series of isolated, blipped shots at different parts of the truth. No overall view of the truth could emerge because each judgment would take place without the benefit of the others. No judgment would be around long enough to be integrated with or informed by others.

I said above that believing a judgment is like writing it down in one's mind – in one's psychology – and this is a useful way to put my point. It's hard to think up a novel all at once. Instead, one writes bits of it at a time so that they can be preserved for later reference and comparison. The book gets written in the process by which the various bits of writing are adjusted to each other and become synthesized.

The drafting and redrafting of a book is similar to the drafting and redrafting of one's total view of what is true; it is ongoing, piecemeal, and gets better with time. For this latter kind of psychological revision to happen though, persons must believe the judgments they make, as this is the equivalent of getting them out onto the "page" of one's mind. Doing so is how judgments can endure so that they can be used in inferences over time, to inform other judgments, and to be shared with other persons. In believing a judgment, one commits to it, invests in it, and so sets oneself up to gain in understanding even if it's false, in the way that one is more on guard for a con if one has been taken in before.[59]

Thus, a belief is an act of thinking, not because believing a judgment provides more evidence for its truth – a judgment does not become more well-supported just because one writes it down or

[59] *See* JOHN STUART MILL, ON LIBERTY 34 (New York: Penguin Books 1985) (discussing how encountering falsehood may strengthen one's own relation to truth).

remembers it later – but because it commits one to integrating the believed content with other judgments.

It's also worth pointing out that a belief is not just a way of thinking more about the content of a judgment, but also a way of understanding the act of judgment itself, by responding to it properly. If one judges that something is true, then one has, in one's own mind, hit on a bit of the truth. And so one should try to retain that moment of success. Believing the judgment is how one does that, because one will keep believing it so long as it seems true.

I have just defended the claim that judging that something is true and then believing it are both steps in thinking. This claim is intuitively true. A judgment is an attempt to think correctly about something, and a belief is, as we say "what one thinks." I defended this claim in the abstract for good measure. I argued that judgments are the building blocks of thinking because they are attempts to have the truth, and beliefs are part of thinking because they allow the building blocks to persist long enough to be parts of theorizing.

With this claim, we can see why punishing a belief amounts to the punishment of a person's thinking. Take, as an example, someone who is curious about whether vaccines are dangerous. He is likely to defer to the experts, but he might not. He might decide to really think about it, and judge that vaccines are dangerous. Assume that this judgment is false, and that it is likely to cause him to act in ways that jeopardize public health.[60] Notwithstanding its falsity and dangerousness, his judgment cannot be punished, because it is an act of thinking. As we say, everyone is entitled to their own opinion, which is to say, the exercise of their rational faculties in the service of a judgment.

[60] *See, e.g.*, Megan Elyse Waterman, *Indorsing Infant Immunity: An Argument for Criminalizing Parents' Refusal to Immunize Their Children*, 51 TULSA L. REV. 153 (2015) (endorsing the assumptions in the text, but not endorsing the criminalization of beliefs).

Of course, under different circumstances, this anti-vaxxer might make a different judgment. Indeed, he might do so in the future when he thinks more or learns new things, but as things stand now, he takes the evidence to support the judgment that vaccines are dangerous. He cannot judge otherwise. Crucially though, the fact that he cannot judge otherwise is not the reason he cannot be punished for it. Instead, he is justified in making this judgment because it is, unavoidably, given how things are for him now, the form that his thinking takes. His judging just is him thinking.

But now, notice, that a similar argument to the one just given will make the punishment of beliefs illegitimate as well. Just as thinking consists of making judgments, it also consists of believing the judgments one makes. The latter is how judgments get put together in a web of judgments that is stronger than the sum of its parts and how one invests oneself into a particular position so that its fate as true or false is meaningful. So, if the interest in thinking protects the judgment that vaccines are dangerous, then it must also protect the belief in that judgment. The latter, just as much as the former, is part of how thinking is sustained and furthered. If it weren't, then we would have to say that the anti-vaxxer does all the thinking he has reason to do by judging that vaccines are dangerous, though he never records that judgment in his mind for further use. That seems untrue. Instead, it seems that persons have reason to believe what they judge, on pain of making the serious rational mistake of not furthering their thinking.

At this point, I have offered a *prima facie* case for thinking that punishing beliefs is wrong. The reason is that their punishment is nothing short of the punishment of a person's thinking (on a subject), which is too valuable to be burdened in that way. This makes sense: what a person believes is just the current form that his thinking takes on its way, in principle, to better thinking. So, punishing the former is a punishment for the latter. Persons need to think, and they think by believing. Hence, beliefs have to be left alone.

To bolster the *prima facie* argument that punishing beliefs wrongly punishes thought, I will now approach it from a second angle; that of showing how punishing a specific belief chills thinking about the topic more broadly, and so punishes the latter too. Scholars have noted that liability for some specific conduct may, in some cases, cause collateral damage to, or "chill" other nearby or related conduct. Sometimes the chilling effect is premised on the fact that innocent conduct may be easily mistaken by a judge or jury for conduct that is illegal.[61]

In other cases though, a chilling effect may be due to the fact that one kind of activity leads, without volition, to another, such that imposing liability for the second provides a reason against engaging in the first. For example, imagine that some jurisdiction makes it a crime to use opioids to satisfy an addiction. In the specific sense that I am using, such a law has a chilling effect because those who wish to comply with it are well-advised not just to refrain from addictively taking opioids (the object of the law's prohibition), but to reduce or refrain from non-addictively taking them too.

The reason is straightforward: one bleeds into the other. If one uses opioids, one risks crossing the invisible line into addiction, and if one becomes addicted, one will very likely violate the law, for two reasons. First, the addiction will motivate one to do what the law prohibits, and second, the addiction will also make one unreasonably insensitive to the punishment one will suffer for doing what the law prohibits. In short then, if one wishes not to use opioids addictively, one will prefer not to use them non-addictively too, because the latter risks the former.

Thinking is, for very different reasons, similar in structure. First, if one thinks, one may be persuaded to think (judge or believe) in a way that is illegal, and second, the punishment one will suffer will not, even if one greatly desires to avoid it, weigh against being persuaded. Thus, if one

[61] *See* Leslie Kendrick, *Speech, Intent, and the Chilling Effect*, 54 WM. & MARY L. REV. 1633 (2013) (pointing out the problems with some kinds of chilling arguments, including those that depend on legal bodies making mistakes).

wishes not to think illegally about some subject, one will prefer not to think about the subject at all, because the latter may lead to the former.

The reason for the first we have already seen. Thinking is fallible and so nothing guarantees that thinking will issue in a true judgment, in line with what is legally required. Thus, someone who starts thinking about vaccines risks judging that they are dangerous, and thereby risks putting themselves at odds with a law that makes that judgment a crime.

The reason for the second is that thinking rightly ignores reasons that are practical. More carefully, one cannot think that something is true in order to secure the practical consequences of thinking that way. As many have put it, the good consequences that will come from thinking that something is true offer the "wrong kind of reason" to think it is true.[62] The right kind of reason to think something is true is, not surprisingly, support for its truth.

A stark example of this wrong-kind-of-reason claim is that one cannot think that the earth is not spherical because an eccentric billionaire has offered to pay one handsomely for thinking that way. This is because though the billionaire's offer makes it good for one to think something is true, it does not provide any support for it actually being true, and without that, one can't think it is true. For one to think something is true, one must be persuaded to judge that it is true.

We can now see why thinking at all about a topic unavoidably risks thinking about it in the way that could be punished, were certain thoughts crimes. Take the vaccine example. If a citizen starts thinking about vaccines at all, then there is a chance that he will judge that they are dangerous, and then believe they are dangerous, and this chance is not reduced by his desire to avoid punishment, since punishment will not be any reason against either the judgment or the belief. Thus,

[62] *See* Hieronymi, *supra* note 27, at 56 (explaining that some reasons are the "wrong kind of reason" for holding a belief, in the ways outlined in this part of the text); Pamela Hieronymi, *The Use of Reasons in Thought (and the Use of Earmarks in Arguments)*, 124 *Ethics* 114 (2013) (same).

whenever a belief is punished, thinking about that topic is chilled in its entirety since the latter may risk producing the punished belief, whereas one is certain to avoid the punished belief by simply forming no beliefs on the subject at all.

An important weakness of chilling arguments is that they can disappear if the world changes in certain ways. For example, punishing addictive opioid use chills non-addictive opioid use only contingently, as things are now. If technology advances sufficiently, it may become possible to ensure that one can avoid an addiction to opioids if one chooses. Were such a breakthrough to come about, then one would not risk addictively using opioids by using them non-addictively.

Similarly, we could suppose that pills were developed and made widely available that prevented persons from forming illegal beliefs, such as that vaccines are dangerous. In this case, it could be claimed, persons could think all they wanted, safe in the knowledge that their thinking would not progress to the point of having an illegal belief. Hence, thinking is chilled by punishing a specific belief only as things are now, contingently, and not necessarily.

Perhaps I must concede this point about the contingency of the argument, but in doing so, let me note that the just-imagined scenario reinforces the first point of this section, which was that beliefs themselves are thinking. After all, I think it is intuitive that a world in which persons must prevent themselves from having beliefs to avoid the chilling effect of laws that punish them is not a world in which thinking is respected, but rather, one in which thinking is under serious attack. A natural explanation for this intuition is the claim I argued for earlier, which is that believing just is (one form of) thinking, so that chemically preventing oneself from believing that vaccines are dangerous is not a way of rendering the criminalization of that belief benign, but of censoring one's own thinking. Hence the availability of the pills would not fundamentally solve anything. One could

either suffer punishment for thinking or else self-lobotomize. Both are unacceptable and so the choice is unacceptable.

Applying the lessons of this section to the resolution author from the introduction is a good way to sum up. The resolution author, like our hypothetical anti-vaxxer, believes that something is true: that all-things-considered, he should attack someone with brass knuckles. Like the anti-vaxxer, this belief is wrong, but the harboring of it is, itself, thinking, and so deserves to be permitted. Punishing it would punish the resolution author for trying to realize the interest he has in thinking his way to the truth. Moreover, punishing a belief in the wisdom of attacking another will chill thinking more generally, since persons may worry that any thinking they do about what is all things considered justified may lead to their judging in favor of something that earns them punishment.

At this point, I conclude my argument against punishing beliefs. In what follows, I extend the argument just given to intentions and answer objections.

IV. Extending and Defending the Argument: Applying it to Intentions and Addressing Objections

A. Applying the Argument to Intentions

I argued just above that persons have an interest in thinking about what is true and that forming beliefs, even mistaken ones, is a crucial part of that enterprise. If the argument is true it shows that the thought crime doctrine is justified in shielding beliefs from punishment. But what about intentions? What justification can be offered for their protection?

In answering this question, I will argue that intentions are like beliefs in being a crucial part of thinking about what is true. It may seem that no argument to this effect could be forthcoming given that intentions are commitments to action rather than representations of what is true. How could they have an integral role to play in thinking about what is true when they are about action?

Though it is true that intentions are "forward looking" in being commitments and motivations to act, they are also "backward looking" in being the enrichments or elaborations of a judgment about what is (truly) worth doing. Thus, persons must be allowed to form intentions in order to complete or round off their thinking about what is true, even though this rounding off takes the form of a commitment to act as the truth requires.

To get some support for this idea, recall that judgments and beliefs are a pair. For example, judging that vaccines are dangerous requires, on pain of making a rational mistake, believing that they are. Judgments and intentions are also a rational pair, but only in the special case where the judgment in question is *practical*, i.e., is a judgment about what to do. If one judges that vaccines are dangerous, this is a judgment that is not, itself, about action. No doubt one might employ it in support of an action, like refusing to get vaccinated, but it is not itself a judgment about what one should do. It's just a judgment about vaccines. Consequently, an intention need not go with it.

But now focus on those judgments that *are*, themselves, about what to do, such as the judgment by skydiver that "right now, considering all the features of my situation, (it is true that) I should jump out of the plane." If one makes this judgment (and does not reverse it) then one should not only believe it, but also intend to jump. Failure to intend would be a rational mistake.[63] This kind of mistake, often called *weakness of the will*, is familiar. Everyone has the experience of judging that some project needs attention, but then forming no intention to work on it, perhaps because it is boring and fun beckons. And certainly, one may judge that now is the moment to jump out of a

[63] This claim is endorsed by many philosophers, and often discussed under the heading of weakness of the will or akrasia. *See, e.g.*, Donald Davidson, *How is Weakness of the Will Possible?, in* ESSAYS ON ACTIONS AND EVENTS 41 (2001) ("what is wrong is that the incontinent man acts, and judges, irrationally, for this is surely what we must say of a man who goes against his own best judgement"); Sergio Tenenbaum, *Akrasia and Irrationality, in* A COMPANION TO THE PHILOSOPHY OF ACTION 274 (Sandis O'Connor ed., 2010) (discussing Davidson). *But see* Nomy Arpaly, *On Acting Rationally Against One's Best Judgment*, 110 ETHICS 488 (2000) (acknowledging, but dissenting from, the "almost universal assumption in contemporary philosophy…that acting against one's best judgment is never an instance of rational action").

plane, but form no intention to jump. My point though is that these failures to intend, however common, are mistakes. We properly regret procrastinating or failing to intend to jump when we judge that it's time.

So why is it a rational mistake not to intend to do what one judges is, all things considered, the thing one should do? For roughly the same reasons that it is a rational mistake not to believe what one judges, which is that the significance of the underlying judgment is amplified or enriched by echoing it in a new form, whether that be a belief or an intention. This kind of enrichment is a bit of thinking so crucial that it is a rational mistake not to carry it out.

Beliefs, I argued, enrich a judgment in three ways: by being a recognition that what is true remains true; by committing oneself more fully to the judgment by accessibly preserving it in one's mind; and, thereby, putting it into conversation with one's other beliefs in a way that permits an overall view of the truth to emerge.

Intentions *further* enrich (practical) judgments in ways that are analogous to those just listed. I will focus on the first two. They do so though not by making a judgment a persisting and accessible part of one's mind as beliefs do, but by transforming a judgment about what one should do into a directive for one's agency or will, in the form of a commitment to make the world over in the way that the judgment requires. What I mean is that one judges *that* "I should jump out of the plane," but the intention that corresponds to this judgment is the intention *to* jump. Intentions thus mediate between thinking and acting; they take, as an input, a judgment about what is true (a truth about what's right or good to do), but they output a commitment to act. It is by transforming a practical judgment into a commitment to act that intentions round off one's thinking about what is true.

Support for this claim, that intentions round off thinking about what is true by transforming it into a directive for one's will, has two main bases.

One is that just in the way that believing a judgment is how one recognizes the general truth that if one thinks something is true, one should keep thinking it is true, so intending, by committing one to act, recognizes the general truth that if one should act, then one should commit to acting.

This recognition of how practical judgments mandate intentions is an increase in understanding of the judgment. It can be compared to the gain in understanding that occurs when one learns more about a concept that figures in something one was told. For instance, one may be told a fact about quarks by a physicist. And one may believe it. But as I argued before, one's belief in the quark fact is imperfect, because though one can repeat it to others and deploy it in certain situations, one doesn't really "get" what a quark is and how it functions in physical theories. Hence, by learning more about quarks, one improves one's understanding of the fact one was told. One enriches one's belief in it. Similarly, by intending to act as one judges (and believes) one should, one understands, more fully, what it means that one *should* act. If one judged that one should act but then sat idly back, one would evince a failure to really *get* the truth that one should act.

Second and relatedly, by engaging one's agency, intentions help one see what a specific judgment means in her particular case. It's one thing to make a general judgment about what should be done (in this kind of situation, the right thing to do is such and such), but it's another to make a judgment about what one must do, here, and now, and still something else to actually commit oneself to doing it. Skydiving is an example. One may readily judge that now is the time to jump, but only by committing oneself to jump does one really appreciate what that underlying judgment really entails.[64] Judgments can be cheap, even if they are true.

Returning to weakness of the will reinforces this point. Many such cases are those in which a person makes the right judgment, but fails to intend. One correctly judges that it's time to start the

[64] *See* Picinali, *supra* note 10 (discussing the potential difficulty of forming an intention).

project, but does not intend to start. But the inverse is possible too. One might make a false judgment about what to do, but then not intend to act that way, and so intend correctly. For instance, one can plausibly interpret Huck Finn as judging, due to the ideology of his society, that he should turn in Jim as a runaway slave, yet he does not intend to turn him in. Instead, he intends to shelter him (and does so), and that is the right thing to intend (and do).[65] What this case suggests is that there is a distinct kind of understanding, had by Huck in virtue of what he intends, about what is right to do. This kind of understanding is infirm on its own, when it does not enrich an already-articulated practical judgment, but it is nonetheless a mode of grasping what is true, and so a part of thinking.

The argument of this section is this: forming intentions is a part of thinking because doing so enriches one's practical beliefs, by drawing out their significance. Therefore, the thought crime doctrine should cover intentions in addition to beliefs. I will now turn to some objections, as addressing them will clarify my view further.

B. Addressing Objections

1. The Argument Proves Too Much: If Intentions Can't Be Punished Then Neither Can Actions, Which is Absurd

One might object that the argument I have given proves too much. I argued that intentions are ways of rounding off one's thinking about what is truly worth doing by enriching the judgments they are based on. But can't it also be argued that actions bring out or enrich the understanding of the intention that motivates them, such that one doesn't really fathom what it is to intend to skydive until one has jumped out of the plane? Put another way: if intentions enrich beliefs owing to rational

[65] Chad Kleist, *Huck Finn the Inverse Akratic: Empathy and Justice,* 12 ETHICAL THEORY AND MORAL PRACTICE 257 (2008) (surveying this case and its implications).

requirement that one intend what one judges there to be reason to do, then won't a parallel claim go through for actions, since it is also a rational requirement to act as one intends? And if actions enrich intentions and intentions enrich practical judgments, then isn't acting a part of thinking, by a kind of transitivity, and thus also, on my telling, immune to punishment?

This objection, in essence, asks me to clarify why thinking can be said to end with an intention, and does not also continue through one's actions. What is the difference that makes a difference between intending to attack with brass knuckles and starting to put on the knuckles, such that the former may not be punished but the second, for all the thought doctrine says, may be?[66]

My answer will be an elaboration of the intuitive idea that actions implement one's thinking about what is truly worth doing, but are not themselves aimed at answering that question. And so acting is not thinking about what is true.[67] This is not to deny that thinking in the very broadest sense of "rational activity" can be predicated of acting. Actions are not just the movements of a machine, but are guided, at all times, by a view of what's worth doing. Nonetheless, action is the execution of thinking about what to do, and not a part of it.

The best support for the intuitive distinction between thinking about what to do and doing it is to re-emphasize the transformative nature of intentions. Intentions are properly formed in response to a judgment *that* some course of action is best or right, but they are not, themselves, about something, but rather commitments *to* do something, such that subsequently motivated actions fulfill that commitment.

---

[66] There might be other reasons, not based in the value of thinking, that militates against punishing the very first actions taken to realize a criminal intention, so-called early attempts.

[67] *See* Jed Rubenfeld, *The Freedom of Imagination*, 112 YALE L.J. 1 (2002) (arguing, similarly, that freedom of imagination protects only the imagination itself, not acting on what one imagines). *See generally*, Laurence Tribe, *The Mystery of Motive, Private and Public: Some Notes Inspired by the Problems of Hate Crimes and Animal Sacrifice*, 1993 SUP. CT. REV. 1 (arguing that the standard *mens rea* requirements for punishment do not punish one for the specific beliefs or ideologies that motivated one to commit a particular crime).

A consequence of this point is that the objection at the beginning of this section relies on a doubtful claim about transitivity. Intentions and beliefs both follow a judgment of the right kind, but actions do not respond directly to a judgment about what is true, but instead, fulfill the commitment to act that is incurred by the intention that governs them. Indeed, part of the justification for forming some intentions in advance ("if he lowballs me, I will walk out of there!") is so that one can act decisively on the intention, without having to confirm that one's action is the right one, which can result in lost time or allow undesired influences to creep in.[68]

Moreover, the fulfillment of an intention is a matter of employing one's practical knowledge, and of exploiting the connections between actions and results so that one can accomplish what one intends by doing what one knows how to do. This kind of knowledge however is not a way of thinking about or enriching one's view of what is true or what one should do, but of taking for granted that one should do as one intends. So though it is true that one does not really understand the intention to, say, dive out of a plane without doing it, the understanding one gets is not one about the worthiness of jumping, but about what it feels like to jump, or of how to jump.

Suppose that I am right that acting is not, itself, thinking about what is true. Still, the objector could press on in a different way. She may concede that punishing action does not punish thinking per se, but she may argue that punishing action chills thinking nonetheless, because nearly every punishable action is the result of some thinking about what to do. So, a law that punishes attacking with brass knuckles will burden judging, all things considered, that one should attack, as well as intending to attack. So, persons will steer away from having such judgments and intentions in order to avoid actually attacking, which will earn them punishment.

[68] MICHAEL BRATMAN, INTENTIONS, PLANS, AND PRACTICAL REASON Ch. 5 (1987).

My response to this chilling version of the objection is that it works by eliding the important difference between revising judgments, beliefs, and intentions on the basis of being given reasons against them, versus steering away from them in a way that is not based on reasons, such as taking a pill that blocks their formation. The first is compatible with thinking; indeed, it *is thinking*. Being given a study that shows that vaccines are safe does not chill the thinking that they are dangerous even though it may lead one to abandon that belief. The second (the pill) is incompatible with thinking; it is an objectionable kind of chilling. Punishing a belief that vaccines are dangerous does not give any reason to judge that they are safe (like the study would) and so encourages persons to self-censor. The reason is that the only way to avoid punishment that is aimed at a conclusion is to stop doing the activity that leads to conclusions, i.e., thinking. In other words, if a possible direct result of thinking is punished, then avoiding that result requires turning off one's thinking, by, e.g., disrupting it with pills or avoiding the relevant subject matter.

Once we see this distinction, we can see that it distinguishes the punishment of an intention and the punishment of an action. An intention is a conclusion of thinking about what to do, and so the only surefire way to avoid an intention is to stop thinking about what to do. Whereas, one can avoid an action by thinking about whether to do it, noticing that it will be punished, and so not doing it.

Take the brass knuckles example. An intention to attack someone with brass knuckles is a conclusion about what to do, and conclusions about what to do issue from thinking about what to do. Consequently, the only surefire way to avoid reaching the conclusion to attack is halt the activity that will produce it: thinking about whether to attack. So, punishing an intention to attack chills thinking about whether to attack.

The act of attacking someone with brass knuckles is different. It is not a conclusion about what to do, but the subject matter of such a conclusion. Hence, if it, the act, is punished, one can still consider whether to attack, since by considering whether to attack, one will consider the punishment as a reason against attack and so not attack. The punishment of an action will "show up" in one's thinking about whether to do it, and so one can avoid the punishment for the action *by thinking about whether to attack*, whereas one can only avoid being punished for the intention to attack by *halting one's thinking about whether to attack*. A different way to put this point is this: because actions are the subject matter of intentions, punishing them provides a new input or consideration to one's thinking about whether to act, whereas punishing the intention does not provide such an input, but only provides a reason to halt one's thinking about whether to act altogether.

I conclude then that punishing actions does not chill thinking, though of course it may make persons think differently, just as scientific evidence does not chill thinking about vaccines, but only changes the conclusion of thinking. Rather, punishing actions allows persons to avoid committing the punished action *by* thinking about the consequences of doing it. That is how deterrence works. It changes the results of a person's thinking; it does not halt it altogether. Consequently, a thinking-based justification for the thought crime doctrine does not rule out the punishment of actions on the grounds that the latter unacceptably chills or deters thought.

2. The Argument Proves Too Little: If Actions Can Be Punished Then So Can Intentions

A different objection picks up where my reply in the preceding section left off. I just claimed that punishing actions does not interfere with thinking because it provides reasons to think differently about acting (if not about the merits of the action apart from its legal consequences), and

not a reason to halt one's thinking altogether. But, a new objection points out, it seems that the same can be said about intentions. That is, punishing intentions can also give one a reason to think differently, which in turn seems to secure the conclusion that punishing intentions is compatible with thinking, even if punishing beliefs is not. In other words, punishing intentions is no different than punishing actions, and so the former must be permissible.

This objection can be illustrated with the following asymmetry.[69] As we saw, one cannot believe something for the practical benefits of so believing. One cannot believe that vaccines are safe because one will be paid for believing it. The reason is that receipt of the money, good though it may be, does not bear on the judgment that vaccines being safe, and so will not support a belief that they are.

Intentions are different. One can intend something for the practical benefits of so intending. One can intend to whistle for money, because the money is a reason one should whistle, given that if one whistles, one will have intended to, and so will have earned the money. Notice though that one can intend as a means to get a benefit so long as the actual action being intended is not so bad as to outweigh the benefit. One would likely not intend to jump off a cliff for money, because the only way to form the intention would be to judge that actually jumping is worth the money, and, let's assume this judgment is false.

If this is right, that one can take into account the consequences of intending to X in judging whether one should X, then it may seem that punishing intentions is really no different than punishing action, viewed from a deterrence perspective. If the punishment for a crime were sufficient, such that a would-be perpetrator would judge against committing it, then the same

[69] Detailed discussion of this asymmetry can be found in the work of Pamela Hieronymi. *See* Hieronymi, *supra* note 27, at 56-8; Hieronymi, *supra* note 62; Pamela Hieronymi, *Two Kinds of Agency, in* MENTAL ACTIONS 138 (Lucy O'Brien and Matthew Soteriou eds., 2009).

punishment, levied against the intention would have the same effect. The would-be perpetrator would see that judging in favor of committing the crime would earn him the punishment, and since, we are assuming, he does not see the crime as worth the punishment, he would not judge in favor of committing it.

But though the deterrent effect may be the same, the route to reaching it is importantly different. I accept that one can intend in order to secure the consequences of so intending, but I contest the assumption that in doing so one is changing one's thinking in response to reasons as opposed to bypassing or manipulating one's thinking through a collateral means. Indeed, I think that the possibility of intending to get a practical benefit is like taking a pill to stifle a criminal belief that I discussed above.[70] One can take the pill to change what one will believe, but the change is not made by thinking about what is true, as it would be if one got new evidence and changed one's mind on that basis. Instead, the change comes from a source, the pill and the decision to take, it, that is alien to thinking about what is true, but rather comes from a practical decision based in the value of not thinking.

Intending so as to reap a benefit for that very intending is similar. Owing to how intentions work, one can pull it off, but in pulling it off, one is gaming oneself. One is not thinking in the proper way, but rather exploiting facts about how thinking works to get a result. A way to put the point is that, owing to how intentions work, persons always have a pill that can forestall their thinking about what to do with them at all times, built in to their psychology.

To elaborate, when one takes the pill to prevent one from believing something, one has the end goal of halting one's thinking. The pill is a means to that end. As we have seen, one might take there to be a reason to adopt this end, say, some prize money for extinguishing the belief, but

[70] *See* text accompanying note 67, *supra*.

nonetheless, in adopting it, one adopts the end of sabotaging or circumventing one's own thinking, through chemical means. In a case where the intention to attack another with brass knuckles is punished, one may choose not to attack so that one will not intend to attack. But here, the decision not to attack is like the intention to take the pill: the reason for it is that doing so will be a means to the end of sabotaging or circumventing one's own thinking. The fact that the "pill" is provided by the structure of one's own agency is immaterial; that makes the sabotage convenient, but does not change the key fact, which is that it is a form of sabotaging or circumventing one's own thinking.

A simple, second way of putting the argument is like this. Say that one was trying to consider whether to attend a party or not. If one weighs the reasons for attending, the result may be an intention to go or not, based in thinking about the merits of going. However, one could decide not to go on the grounds that deciding not to go would make it true that one did not have an intention to go. If one decides this way, then one has not made a decision about whether it would be worth going and so one's lack of an intention does not reflect a judgment that it's not worth going. Instead, the lack of one's intention is what one values in this case, for itself, and not as a consequence or reflection of one's thinking that the party is not worth doing. This way of thinking about the party is distortive, since it elides the question of whether it's worth going to the party and instead answers the question about going to the party on the basis of what that answer will cause in oneself.

V. Conclusion

I have addressed the question: what is the justification for the thought crime doctrine, which forbids, at least, the punishment of beliefs and intentions? My answer relies on claims about the value and structure of thinking.

I argued that thinking is valuable because it is how one engages with ideas in an effort to understand them. This kind of engagement, the daily doing of it, is a large part of what being a person is about, as it is through this engagement that we grow and succeed at almost all else.

I also argued that thinking is structured as a cumulative, progressing activity that works by, or through, mistakes as much as successes. One practices something just by earnestly and self-consciously doing it, with the goal of getting better, and thinking works the same way. One does it just by earnestly and self-consciously trying to understand what is true. Since one does this by believing and intending, beliefs and intentions are sacrosanct, whether correct or not. To punish them is to punish thinking itself, which is absolutely central to developing and living as a person, and so forbidden.

This view has significant advantages over the other two I considered primarily by showing how a seemingly anodyne idea – that thoughts, whether beliefs or intentions, are bits of thinking – can be worked into a full justification for the thought crime doctrine.

Duff assumes that only completed crimes can be punished. But this principle is unmotivated. It does not distinguish thoughts and criminal attempts. Also, it has the unintuitive consequence that criminal attempts cannot be punished. By contrast, the view I offer gives a principled and distinct reason for why thoughts should not be punished: because they are, unlike attempts, part of thinking about what is true.

The view I offered also avoids my criticisms of Mendlow because it does not work by assuming a moral conclusion regarding thoughts; i.e., that they may not be disrupted. Instead, it operates by showing that thinking is fundamentally valuable and that it cannot proceed without generating thoughts, whether mistaken or not. Moreover, the justification I offer does not depend on the contestable link between disruption and punishment that Mendlow puts forward in the

enforceability constraint. Instead, it shows how punishing thoughts, and disrupting them too for that matter, are both wrongful interferences with thinking. By showing how thinking is wrong, the view I have defended has the potential to generalize, in a way that Mendlow's does not, to other kinds of interference, such as taxes on thought.[71]

Given these advantages, a thinking-based theory is best positioned to justify the full extent of the thought crime doctrine. The doctrine covers beliefs and intentions, and because both are aspects of thinking, they earn protection on that basis.

[71] *See* text accompanying note 50, *supra*.

Chapter 4: Self-Incrimination

I. Introduction

In previous chapters, I argued that persons may not be held responsible for their thoughts as such, whether by curious neighbors or by the community, embodied in a legal regime. However, many thoughts do not remain just thoughts. They make an entrance into the world through actions, and actions are not fragile in the way that makes thoughts generally off limits to others. Indeed, actions are sound targets for practices of holding responsible. Still, holding persons responsible for their actions may endanger thoughts in some cases, such as when a person is held responsible for an action *by* holding her responsible for her memories of doing it. This can happen dramatically in a legal context, if someone is forced to be a witness against himself.

For at least the last fifty years, commentators have refuted proposed justifications for the Fifth Amendment privilege against self-incrimination.[1] The core of the privilege is the right not to testify in one's own criminal trial, but it also includes the right not to respond to other kinds of government questioning when doing so would help establish one's criminal guilt. In some cases, the privilege also permits one to ignore government document requests, when answering the request would itself provide incriminating evidence.[2] The canonical rationales for the privilege offered by

[1] U.S. CONST. amend. V ("No person…shall be compelled in any criminal case to be a witness against himself"); Robert McKay, *Self-Incrimination and the New Privacy*, 1967 SUP. CT. REV. 193 (1967); David Dolinko, *Is there a Rationale for the Privilege Against Self-Incrimination?*, 33 UCLA L. REV. 1063 (1986); Akhil Amar & Renee Lettow, *Fifth Amendment First Principles: The Self-Incrimination Clause*, 93 MICH. L. REV. 857 (1995).

[2] *See* United States v. Hubbell, 530 U.S. 27, 36 (2000) (discussing this doctrine at length and holding that compelling the production of broad categories of documents would violate the privilege against self-incrimination because their production would confirm their existence and the defendant's possession of them).

Justice Goldberg in *Murphy v. Waterfront Commission* have come in for especially sustained criticism over the years, with some justice.[3]

Justice Goldberg claimed that the privilege saves defendants from having to make an agonizing three-way choice between contempt, perjury, or "self-accusation" – but this begs the question.[4] Why should the defendant be allowed to avoid testifying, since it won't be incriminating or self-accusatory unless he is guilty? Justice Goldberg claimed that the privilege evinces a preference for an accusatorial over an inquisitorial system, which it does, but why indulge it, and what necessary role is played in such a system by a right against self-incrimination?[5] He thought the privilege preserved a fair "state-individual balance," but this is still conclusory.[6] On its face, the privilege seems to disadvantage the state in an ad hoc and substantial way. Justice Goldberg also suggested that "self-deprecatory" testimony is not to be trusted, but categorically so?[7] It seems the probative value of such testimony could be substantial and its risks ameliorated without complete exclusion.

Criticism of these rationales and their subsequent elaborations has been broadly persuasive – so much so that it has seemed to some that it is time to abandon the doctrine as unjustified. In a 2008 symposium on self-incrimination, one leading commentator stated that the outstanding question regarding the privilege is not "'what is … the best, theory of self-incrimination,' because it

[3] Murphy v. Waterfront Comm'n of NY Harbor, 378 U.S. 52, 55 (1964) (holding that persons may exercise the privilege against self-incrimination even if the testimony they would be compelled to give would only expose them to prosecution in another jurisdiction).

[4] *See* Amar & Lettow, *supra* note 1, at 890; Michael Green, *The Privilege's Last Stand: The Privilege Against Self-Incrimination and the Right to Rebel Against the State*, 65 BROOK. L. REV. 627, 630-1 (1999); Dolinko, *supra* note 1, at 1092; Stephen Schulhofer, *Some Kind Words for the Privilege against Self-Incrimination*, 26 VAL. U. L. REV. 311, 318 (1991).

[5] *See* Dolinko, *supra* note 1, at 1088-89; Amar & Lettow, *supra* note 1, at 893.

[6] *See* Amar & Lettow, *supra* note 1, at 891; Schulhofer, *supra* note 4, at 317.

[7] *See* Amar & Lettow, *supra* note 1 (arguing that some limitations on the self-incrimination privilege would be consistent with securing accurate testimony); Schulhofer, *supra* note 4, at 319.

is plain that there is and will not be any good ones of the standard sort" but rather "the puzzling persistence of the futile effort at standard theorizing."[8]

I am less pessimistic. A good, principled justification for the privilege has been with us from the start of the modern jurisprudential era, in *Murphy*: namely, that it respects "the right of each individual to a private enclave where he may lead a private life."[9] Admittedly, the wording does not provoke much sympathy – does the defendant's testimonial cooperation with the investigation of serious lawbreaking deny him a private life?

The idea is expressed better in *Couch v. United States* a few years later, which focused specifically on the privacy of the mind – of an "inner sanctum of…feeling and thought."[10] Some motivation for this way of viewing the privilege comes from the possibility, currently remote but getting closer, of analyzing a defendant's brain to gain evidence about his wrongdoing.[11] If such capabilities became sufficiently accurate, many rationales for the privilege would become obsolete. Nonetheless, one may feel that using mindreading against a criminal defendant is something that the privilege squarely condemns. A plausible explanation is that mindreading violates, and the privilege against self-incrimination protects, mental privacy.

It's not as if (mental) privacy has escaped scrutiny in discussions of the privilege against self-incrimination. It has been repeatedly rejected by commentators along with the other *Murphy*

[8] Ronald J. Allen, *Theorizing About Self-Incrimination*, 30 CARDOZO L. REV. 729, 730 (2008). *See also*, Kenworthey Bilz, *Self-Incrimination Doctrine is Dead; Long Live Self-Incrimination Doctrine: Confessions, Scientific Evidence, and the Anxieties of the Liberal State*, 30 CARDOZO L. REV. 807, 841 (2008) ("If a few hundred years of scholarship and interpretation haven't settled the confusion [about the privilege against self-incrimination], it's hard to imagine anything that will.").

[9] Murphy v. Waterfront Comm'n of NY Harbor, 378 U.S. 52, 55 (1964).

[10] Couch v. United States, 409 U.S. 322, 327 (1973) (the privilege protects an "inner sanctum of individual feeling and thought.").

[11] *See* Henry Greely, *Neuroscience, Mind-Reading, and the Courts: the Example of Pain*, 18 J. HEALTH CARE L. & POL'Y 171 (2015) (surveying current scientific capabilities).

rationales and their spin-offs.[12] However, the debate has been fairly lopsided. More energy has been put into criticizing the privacy idea than into developing it, with the exceptions largely consisting of three early articles by Robert Gerstein (1970, 1971, and 1979 respectively).[13] Since then, the privacy view has fallen into desuetude.[14] In the nineties and more recently, the view is often refuted quickly, usually with the observation that it does not square with judicial reasoning,[15] or that it is inconsistent with the exact contours of doctrine, such as the permissibly compelled testimony of civil witnesses or immunized witnesses.[16] The persistence of these criticisms is justified, given that the normative underpinnings of privacy and its contours are heavily disputed. Still, attention to the latter will provide resources to address the former, at least so I claim: in this chapter I argue that the privacy rationale can be updated and made good.

My *prima facie* case for the privacy rationale has much in common with Gerstein's somewhat neglected third article on self-incrimination, in which he argues that "an individual ought to be autonomous in his efforts to come to terms in his own conscience with accusations of wrongdoing against him."[17] My way of putting the idea can be encapsulated like this: self-incrimination is a morally risky way of enforcing criminal norms because it enlists a kind of thinking – reflection on

---

[12] *See* Allen, *supra* note 8, at 734; Dolinko, *supra* note 1, at 1108-1137; Amar & Lettow, *supra* note 1, at 890-1; Green, *supra* note 4, at 659; Nina Farahany, *Incriminating Thoughts*, 64 STAN. L. REV. 351, 360-2 (2012) (rejecting the privacy rationale but also finding it normatively attractive); Schulhofer, *supra* note 4, at 319-320.

[13] Robert Gerstein, *Privacy and Self-Incrimination*, 80 ETHICS 87 (1970) [hereinafter *Privacy*]; Robert Gerstein, *Punishment and Self-Incrimination*, 16 AM. J. JURIS. 84 (1971); Robert Gerstein, *The Demise of Boyd: Self-Incrimination and Private Papers in the Burger Court*, 27 UCLA L. REV. 343 (1979) [hereinafter *Demise*]. David Dolinko raised several objections to the earlier article. Dolinko, *supra* note 1 at 1122-1137. Stephen Shulhofer discusses Gerstein's earlier article but only briefly. Schulhofer, *supra* note 4, at 320.

[14] *E.g.,* Peter Arenella, *Schmerber and the Privilege Against Self-Incrimination: a Reappraisal*, 20 AM. CRIM. L. REV. 31 (1982) (avowing a mental privacy view, but not defending its normative basis).

[15] *E.g.,* Farahany, *supra* note 12, at 360-1, Schulhofer, *supra* note 4, at 320 (calling Gerstein's 1971 article "sophisticated" but not discussing it at length); Kiel Brennan-Marquez, *A Modest Defense of Mind Reading*, 15 YALE J. OF L. AND TECH. 214, 261 (2013) (criticizing Gerstein's arguments on the grounds that they are not incorporated into doctrine).

[16] Amar & Lettow, *supra* note 1, at 890, Allen, *supra* note 8, at 734; Kiel Brennan-Marquez, *supra* note 15, at 264 n.193; Akhil Amar, *Against Exclusion (Except to Protect Truth or Prevent Privacy Violations)*, 20 HARV. J.L. & PUB. POL'Y 457 (1997); William Stuntz, *Self-Incrimination and Excuse*, 88 COLUM. L. REV. 1227, 1232-34 (1988).

[17] Gerstein, *Demise*, *supra* note 13, at 344.

the moral quality of one's actions – that persons have an interest in performing, to some extent, unburdened. The argument does not depend on the criminal law being unjust or the defendant being falsely accused. Mistakes and miscarriages in the content and application of laws are pervasive in real life, but I assume throughout that they are absent to avoid distracting from my point, which is that even when a fair law is fairly enforced against a true violator of it, there is an outstanding moral concern with compelling the violator to incriminate himself. This assumption is only for the sake of focusing the discussion. By making it, I am not contending that the privilege against self-incrimination has no relevance to the innocent. It may. I discuss this issue briefly in the conclusion.

The moral concern unpacks into two main premises. First, a principle: persons have an interest in thinking, one species of which is thinking about which actions they wish to take responsibility for, and so hold out for recognition as a part of their biography, or self. For example, in writing a memoir, one takes responsibility for different aspects of one's life, by weaving them into a narrative that one wishes others to accept. The starting point for this kind of thinking is a memory of what one did, which permits further reflection, such as evaluating the reasons one had for doing what one did, and perhaps, in concert with such evaluation, adopting various attitudes toward it, such as regret, guilt, remorse, shame, or embarrassment.

The second premise is that the circumstances of criminal prosecution place great pressure on a person's moral self-reckoning, by confronting them with, and so encouraging them to unreflectively adopt, the majority's view of what they have done and its gravity. Though a defendant may be held responsible for breaking the law, this holding responsible cannot be so zealous as to deny the defendant a chance to take responsibility for what he has done, which is a matter of engaging, in a personal and unmediated way, with the reasons for the law and one's violation of it.

After defending this argument, I illustrate its advantages and try to fend off some of its seeming weaknesses. The greatest challenge in this area is trying to accommodate the apparently well-justified limitations on the privilege. It is inapplicable to civil cases or to immunized witnesses. Hence, there are needed refinements to the rationale I offer. The refinements will illustrate the ways in which the principle I isolate cannot be applied blindly to legal doctrine but instead interacts with other values and doctrines to produce a fair regime of post-wrongdoing reflection.

The chapter is organized as follows. First, I discuss the way in which persons take responsibility for actions that are not speech. Next, I make a *prima facie* case for the privacy rationale by defending the two premises listed just above. In doing so, I discuss the argument's similarities, but also differences, from the one offered by Robert Gerstein in his insightful 1979 article. Third, I defend it against some common arguments against privacy being the basis of the privilege against self-incrimination.

II. A Mental Privacy Theory of the Privilege Against Self-Incrimination

A. Taking Responsibility for Non-Speech Actions, Such as Criminal Wrongs

In this section, I explain how non-speech actions like robbing a bank or attacking someone with brass knuckles differ from thoughts with regard to when persons may be held responsible for them and when, and how, they should take responsibility for them. I end up with the following picture: the commission of a crime is the public's business, and hence, someone who commits a crime may be held responsible for doing so. Moreover, someone who commits a crime should take responsibility for it, which they do by avowing its commission to the proper authorities. However, this duty is not, as I argue in the next section, one that can be legally mandated.

In Chapter 1, I argued that a person may not be held responsible by others for the content of their thoughts, but that they may be held responsible for their (non-mental) actions that are based on that content.[18] One may be held responsible for them. For example, a belief with the content "global warming is caused by humans" is the basis for the action of sincerely asserting that same content to someone. Alternatively, an intention with the content "cut in line" may be the basis for cutting in line. In both cases, one may not be held responsible for the thought, but one may be held responsible for the action that is based on it. Actions, just as actions, are suitable items for which to be held responsible.

But there is also a distinction between these two actions, that must now be faced in more detail for the purposes of this chapter. The distinction is that speech is by nature suited for communicating things, but non-speech is not. This is not to say that non-speech actions cannot be used to communicate. They often can be, but in a derivative way. A consequence of this distinction is that speech is by nature suited to taking responsibility for one's thoughts and thereby securing a mutual understanding, whereas non-speech is only sometimes suited to the same.[19] Recall that taking responsibility is matter of inviting another to link one to some content by making that linkage sufficiently manifest. Consequently, a reliably clear way of issuing such an invitation is by speaking the content one wishes to be linked with, in one's own name, to a target recipient. Doing this allows the other to succeed, in turn, in holding one responsible for one's thoughts, both by providing strong evidence that one wishes to be held responsible for them, but also by providing their relevant

[18] Mental actions are things like imagining a scenario or rehearsing some lines of a play. My focus is exclusively on non-mental actions because crimes are invariably non-mental actions. I drop the non-mental qualifier in what follows and just talk about action.

[19] *See* SEANA VALENTINE SHIFFRIN, SPEECH MATTERS 112-5 (2014) (discussing the communicative clarity of speech as compared to less speech-like actions).

content in a highly explicit form. Put simply, speech sets up a robust mutual understanding about what one thinks.

Actions besides speech can do this too. Some actions, like a thumbs up (to communicate acknowledgment and affirmation), or cupping one's hand to one's ear and straining as if listening (to communicate a difficult in hearing) are word-like and so can create a mutual understanding. Other actions may take on a clear meaning with help from the surrounding situation. Extending one's hand on a dance floor communicates an invitation to dance. In other cases, the communication that is achieved by non-speech action is attenuated in various ways, and hence the corresponding invitation loses its clarity. For instance, burning a flag certainly communicates something, roughly, disapproval of a regime or its policies, but the nature of the criticism and its exact target may be unclear – wars in the Middle East? All wars? All wars in the Middle East for oil?

Still other actions fail to initiate or make possible a mutual understanding because they are not intended to be picked up by anyone, even if the content that they manifest is unambiguous. Hearing someone mutter something under his breath, when one can tell he thought no one was around, may make some content clear, but obviously is not meant to invite anyone to hold him responsible for it. An even more extreme example is an accident. Plainly, one does not invite others to hold one responsible for what one does not mean to do.

The relevance of these points about taking responsibility can be seen in their application to the shopper who cuts in line. Conceivably, someone may cut in line so obviously and with such a knowing smirk, that he invites others to hold him responsible for doing so.[20] In such a case, an onlooker might hold him responsible for cutting just by seeing him do it, with the result being an

[20] Obvious wrongdoing that seeks the recognition of those it wrongs is an aggravated kind of wrongdoing that seeks to mock or intimidate the victims.

immediate mutual understanding *that* he cut. This mutual understanding is here constituted just by an action done in a particular way and the witnessing of it. Alternatively, someone may cut, realize that he did not check whether anyone was already in line, turn around, see someone there and say "sorry," and then leave the line. In this case, there is a mutual understanding that he cut, as evidenced by the fact that there is no doubt *what* he was sorry for, because it – the cutting – was sufficiently salient to both parties.

However, cutting need not communicate anything to anyone. We could imagine someone cutting unassumingly without anyone noticing. In this case, cutting is just something that gets done, and it does not invite anyone to hold the cutter responsible for doing so.[21] Alternatively, we could imagine that the cutter was not trying to cut, but did so negligently (he didn't see the obvious line). Others in line let it go, not saying anything, and time passes. Later, he turns and says "sorry." It may not be clear what he is apologizing for, and so some speech may be needed to clarify.

The point is this: (non-mental) actions are ways of putting the content of a thought out into the world. Speech and non-speech are, as actions, equally ways of accomplishing this, and hence persons may generally be held responsible for both what they say and do. However, speech is a special kind of action because it is built for taking responsibility for content, and so for creating a mutual understanding. Non-speech action is not built for taking responsibility, but it can be pressed into service for that end in some circumstances.

Importantly, when non-speech action fails to create a mutual understanding, for any number of reasons including that one performed it covertly, one can still take responsibility for it, by using speech to avow it, and thereby provide a kind of explanatory coda or explicit acknowledgment. This

[21] One may talk to oneself while alone, and clearly this does not invite anyone to hold one responsible for what one says. However, speech is for communicating even if it is not used that way, but cutting is not. It's just for cutting, even if it may be used to communicate.

coda can be provided before or after the action, to different effect.[22] I have focused on the "after" case, in which one takes responsibility for what is already done by saying one did it. There are two steps here: the doing of a non-speech action makes it permissible for one to be held responsible for it, and by avowing it later one takes responsibility for it and so makes possible a mutual understanding that one did it.

Persons take responsibility for what they do to all sort of different audiences. In many interpersonal cases, a request to take responsibility is justified and a refusal to do so may justify the other in drawing negative inferences. If one friend asks another where he was during her piano recital and gets no answer, she may infer he had no good reason to miss it. By contrast, if someone commits a crime and refuses to testify at her trial, the privilege against self-incrimination prevents the jury from inferring that she is guilty. I think a reason for this asymmetry is that an interpersonal relationship is partly constituted by a certain kind of joint deliberation, so that by refusing this deliberation, one injures the relationship, and one has no right to be in a relationship with another person when one refuses to participate in it. However, the state cannot and may not dissociate itself from someone who refuses to take responsibility for what they have done. They have to vindicate the rights of the community and so have to go about scrupulously creating the evidentiary basis to do so.

I have discussed how persons are held responsible and take responsibility for actions. I now turn to the issue of when these two activities are appropriate. Roughly, their appropriateness is determined by the kind of deliberation that is called for, regarding the action.

[22] One can take responsibility for an action one has not yet done, by stating that one will do it. Stating what one will do is, ordinarily, a making of a promise and gives one an obligation to do it.

For example, consider Kyle, who gets a bonus at work, but squanders it on a frivolous purchase. May he be held responsible for the purchase? Need he take responsibility for it? No to both, with regard to, say, his neighbor. This is because his neighbor is not a co-deliberator with him about his finances. It's none of his business. Because it's not his business, he should not be trying to find out how Kyle's bonus got spent, as finding out would hold him responsible for squandering it. Correspondingly, Kyle need not take responsibility for his frivolous purchase to his neighbor. He doesn't have to tell him about it.

Kyle's partner, Kylie, is a different matter. She *does* have a claim to be an equal deliberator on matters of household finances. The bonus is her business. Consequently, she may hold Kyle responsible for squandering it, which in turn gives her some permission to seek that information through investigation. Correspondingly, Kyle should take responsibility for squandering the bonus, and his failure to bring it to Kylie's attention compounds his mistake in wasting it.[23] Whether Kylie confronts Kyle, or he comes forward first, a dialogical process will ensue in which the parties will, assuming they are both committed to the relationship, come to a mutual understanding of what Kyle did (perhaps he denies it for a day or two and then comes around, or perhaps Kylie finds out and does not confront him to see if he comes forward first), which will, in all likelihood, provoke a conversation about its meaning or significance going forward, which may in turn require Kyle to apologize, or take other steps.

Kyle and Kylie's case provides a way to start thinking about politics. The nature of their relationship and the activities that it encompasses makes it that certain things are each other's business, but not the business of their neighbors, e.g., their finances. The same can be said of

[23] Jeffrey Helmreich, More than Words: Stances as an Alternative Model for Apology, Forgiveness and Similar Speech Acts (2013) (unpublished Ph.D. dissertation, University of California, Los Angeles) (on file with University of California eScholarship) (available at, https://escholarship.org/uc/item/3z03x8qv) (arguing that a failure to apologize leaves a wrong unaddressed, and that this failure is a distinct kind of wrong in itself).

citizens. Their relationship and the activity constituting it – politics – give rise to the public's business in the form of laws. Crucially, there are principled restrictions on what is eligible to be the public's business. What citizens think may not be the public's business. And speech, though it is action, is so suffused with thought that it may not be the public's business either, at least in the main.

When it comes to non-speech, non-mental actions though, I doubt there are any such categorical restrictions. After all, such actions alter the world, and the world is shared by those who live in it, and so there may be reason for the public to make such actions its business, through law. For instance, watering the front yard may be one's private business when water is plentiful, whereas drought conditions may warrant water use restrictions. Thus, whether actions are public or private, or which spaces are public or private is a matter, in principle, of public deliberation and decision.

I do not mean that there are no stable or weighty reasons that bear on whether something should be private or public. I think there are reasons, for example, why the home has the status that it does in, say, search and seizure jurisprudence.[24] The home is a place to which one may retreat from various kinds of expectations so as to lower one's guard, experiment, think, and interact with particular others in one's own style.[25] Consequently, it would be wrongful to require persons to live in perfectly transparent structures or to allow the authorities to breeze in anytime they please.[26]

---

[24] *See, e.g.*, United States v. Karo, 468 U.S. 705, 714 (1984) ("private residences are places in which the individual normally expects privacy free of governmental intrusion not authorized by a warrant, and that expectation is plainly one that society is prepared to recognize as justifiable."); James Q. Whitman, *The Two Western Cultures of Privacy: Dignity Versus Liberty*, 113 YALE L.J. 1151, 1211 (2004). *But see* Stephanie M. Stern, *The Inviolate Home: Housing Exceptionalism in the Fourth Amendment*, 95 CORNELL L. REV. 905, 906 (2010) (arguing the home is excessively protected in search and seizure jurisprudence).

[25] For a different perspective see Chris Essert, *Property and Homelessness*, 44 PHIL. AND PUB. AFF. 266 (2016) (discussing property rights in a home as a necessary kind of freedom).

[26] *See* WE LIVE IN PUBLIC (Interloper Films, 2009) (depicting a (ultimately dystopian) community experiment in which residents lived with each other in complete transparency).

Instead, my point is that these reasons for the privacy of the home are properly assessed and approved by the public, since the public interest is, in principle, implicated in every decision about what actions and spaces are made private. After all, a decision to make something private will put it out of reach of oversight and control. Put another way, the privacy of the home is, though well-supported, unlike the privacy of the mind in that it is not an external limitation on the public, but rather constitutes the proper recognition, *by the public*, of reasons to moderate, but not absolutely curtail, its own investigatory activities.

A simple example will illustrate the arguments of this section and their relevance to self-incrimination. Take the criminalized wrong of attacking someone with brass knuckles. The law against doing this is the announced judgment of the public that such attacks are its business. By violating this law, a defendant contravenes this judgment, and so makes the legal system justified in holding him responsible, where this is, in the first instance, a matter of establishing whether he did it. Correspondingly, the defendant should come forward and take responsibility for violating the public's judgment, which he does by admitting to the relevant legal officials that he committed the crime. If he does not admit that he committed the crime, then the government will have to proceed unilaterally, offering its own version of events to a neutral party. If its version of events is accepted, it has a basis to proceed with more, such as punishment or other remedies.

I have emphasized the way in which avowing one's criminal action with speech takes responsibility for the crime itself because it plays the theoretically important role of differentiating the testimony of the defendant from that of other witnesses. Other witnesses may take responsibility for the truth of the statements that they make, and some of those statements may concern actions, such as actions they took, or that the defendant took, but by taking responsibility for the truth of

these statements, they are not thereby taking responsibility for the conduct that is at issue in the trial, and so, not taking responsibility for it in the way that the defendant's testimony does.

Without this distinction – between taking responsibility for the truth of a statement about what one did, and taking responsibility for what one did – it's hard to make sense of the privilege against self-incrimination, because it will seem that defendant is similarly situated to ordinary witnesses, who must testify to the truth of what they know.

For instance, Michael Pardo acknowledges the need to distinguish the defendant's testimony from that of other witnesses, and he correctly conceptualizes testimony as being the taking of responsibility for the truth of what one asserts (he calls this assuming epistemic authority for one's testimony).[27] However, he stops there. He does not claim that by taking responsibility for statements about what one did, one thereby takes responsibility for the actions themselves. Hence, he has to look elsewhere for a way to distinguish the defendant from other witnesses.

This he does epistemically, with the presumption of innocence.[28] He argues that compelling a defendant to assume epistemic authority for incriminating statements (or for exculpatory statements that can be attacked) is "perverse" because it would allow the government to pass the burden imposed by the presumption of innocence to the defendant, by requiring him to put together a case against himself.[29]

I don't agree. The government does not dodge the epistemic burden imposed by the presumption of innocence by requiring ordinary witnesses to testify.[30] Rather, the government *fulfills* its burden by requiring them to assume epistemic authority for what they know. So, why not the

[27] Michael Pardo, *Self-Incrimination and the Epistemology of Testimony*, 30 CARDOZO L. REV. 1023 (2008).
[28] *Id.* at 1042.
[29] *Id.* at 1044.
[30] *See* Green, *supra* note 4, at 652-3.

defendant? Asking the defendant to testify does not ask the defendant to take responsibility for building the government's case for it. If the defendant's testimony is inculpatory and true, it may still be disbelieved by the jury. And if the defendant's testimony is exculpatory but false, the government must work to show that it is indeed false. Either way, the government would not shift any of its epistemic burdens to the defendant by making him testify, but would only try to satisfy those burdens by requiring the defendant to take epistemic authority for his testimony.

It is this latter assumption of authority that I think is troublesome, not epistemically, but morally. I argue this next.

B. A Consideration Against Compelled Self-Incrimination: The Mental Privacy Dangers of Confessional Situations

In this section, I present a *prima facie* argument for the privilege against self-incrimination. The major premise of the argument is that memories about a crime that one has committed constitute an important kind of thinking; namely, self-reflection about whether to take responsibility for it. The minor premise is that holding persons responsible for these memories (by revealing them, whether through compelled testimony or mind reading technology) as a way of holding the person responsible for the actions they remember distorts and corrupts self-reflection.

In defending these premises, I will occasionally make reference to Robert Gerstein's 1979 article on self-incrimination, which has not garnered much critical discussion.[31] In that article, he develops his earlier claims about the relationship between self-incrimination, privacy, and the repair of wrongdoing, in a productive direction. Specifically, he dispenses with the idea that the privilege against self-incrimination respects the privacy of a discrete judgment about oneself (a judgment of

[31] Michael Pardo, *Disentangling the Fourth Amendment and the Self-Incrimination Clause*, 90 IOWA L. REV. 1857, 1862-3 (2005) (mentioning Gerstein the course of discussing privacy rationales for the privilege); Dolinko, *supra* note 1, at 1122-1137 (almost exclusively discussing Gerstein's first, 1971 article).

self-condemnation)[32] and instead sees it as a protection of a person's reflective moral thinking about what they have done.[33] This aspect of his view, its major premise, seems to me to be an important step forward and parallels the major premise I defend just below. However, his defense of the requisite minor premise, that self-incrimination damages this kind of thinking, is quite brief, and so leaves some work for even a sympathetic reader. I try to remedy the lacuna.

1. Self-Reflection is Valuable Form of Thinking

My first premise, that thinking about taking responsibility for action is valuable, is prefigured by Gerstein.[34] It is also a specific instance of a claim I argued for in Chapter 1, which is that thinking is valuable, and that its value is realized in stages, with a personal engagement with reasons being a necessary and preparatory step to thinking with others. Here, I assume those arguments as background and focus on how they apply to self-reflection about what one has done. This kind of thinking can be summed up as a retrospective encounter with the moral reasons, *inter alia*, for one's actions, judged afresh in light of the effects of the action on others as well as their reactions to it.

A good way to overview the kind of argument I will be giving is to show how retrospective thinking about a crime one has committed is the continuation of prospective thinking about whether to commit it. This results in a high-altitude picture of the legal norms that protect the mind.[35]

Recall my example from the previous chapter: an intention to commit the criminalized wrong of attacking someone with brass knuckles. Such an intention is the mental *antecedent* to the

---

[32] This idea is in Gerstein's first article. Gerstein, *Privacy*, *supra* note 13 at 90.

[33] Gerstein, *Demise*, *supra* note 13, at 396 ("…we must be free to come to terms with our consciences in our own time, and to choose either to condemn ourselves by confessing or to stand aside in dignified silence from the government's effort to prove our guilt…").

[34] *Id.* at 348 ("[W]e have the capacity to make our own judgment as to the rightness or wrongness of our conduct and, if we come to see ourselves as guilty, the capacity to reconcile ourselves with ourselves and with others through acknowledgement of guilt and acceptance of responsibility. In exercising this capacity we come to know ourselves as moral beings and to develop standards of conduct.").

[35] H. Richard Uviller, *Evidence from the Mind of the Criminal Suspect: A Reconsideration of the Current Rules of Access and Restraint*, 87 COLUM. L. REV. 1137 (1987) (engaged in a similar project).

actual commission of that crime, and for that reason, it seems to be something that the legal system, given its interest in *crime prevention*, might be warranted in trying to discover and suppress.[36] Yet, I argued, it is not so warranted, and instead, the legal system must confine itself to what persons do toward realizing such criminal intentions (the subject matter of the law of attempts). The reason is that intentions, even those with a wrongful aim, are a stage in thinking and so should not be something for which a person is held responsible.

Now consider the memory of having carried out the attack with brass knuckles. It is the mental *consequent* of the commission of that crime, and for that reason, it seems to be something that the legal system, given its interest in *crime prosecution*, is warranted in discovering and using it. Yet the privilege against self-incrimination forbids this use, and confines its interest in the defendant to the state of his body, and to forbidding actions that would constitute interference with the discovery and use of real evidence, or testimonial evidence from others.

The reason for the parallel legal anxiety about[37] mindreading to prevent crime or to prosecute it is that the thoughts being "read" – intentions aimed at future wrongdoing and memories of past wrongdoing – are both perspectives, one prospective and the other retrospective, on the same fundamental question: what is worth doing? Persons must be able to think about this question, and hence, incursions by others, including the government, into the thoughts that constitute such thinking is damaging. The burden of this section is to make this claim more concrete, as applied to retrospective thinking about what one has done.

---

[36] In the previous chapter, I focused on the criminalization of an intention to commit a crime. However, it would also clearly be a thought crime to criminalize the failure to disclose one's criminal intentions to the police. Would such a law offend against the thought crime doctrine or against the Fifth Amendment privilege against self-incrimination? I think it straddles both, suggesting the deep connection that underlies them.

[37] This anxiety appears in fiction and in scholarship. *See* Bilz, *supra* note 8. In fiction, mind reading that is used to prevent or prosecute crime is viewed with suspicion. *See, e.g.*, PHILIP K. DICK, MINORITY REPORT AND OTHER CLASSIC STORIES (Citadel 2016) (1956) (crime prevention), ALFRED BESTER, THE DEMOLISHED MAN (1951) (telepaths that can probe thoughts may not be used to establish a person's criminal guilt).

As an analogy, consider legislative thinking and judicial thinking in the U.S. constitutional system. A legislature considers a law prospectively and abstractly. Attempting to alter the past through an *ex post facto* law or singling out persons through a bill of attainder are forbidden.[38] Courts however will be called on to evaluate the same law in a different posture, i.e., in the course of applying it to a specific factual situation. This difference, between prospection and retrospection, is a difference in the kind of thinking being done, and hence the combination of legislative action with judicial review can claim some advantage, in that it is a way of thinking thoroughly about political issues.

The life cycle of our own actions – our personal "legislation" – is not so different. Before acting, we ask ourselves the wide-open question "what should I do?" and then form a corresponding intention. The intention is one's mental answer to that question, which, when executed, embodies the answer in the world, where one can be held responsible for it. However, an action is not a conclusive answer to the question "what should I do?" so much as a way of inaugurating a different way of trying to answer it, now from a position of having given an answer. To restate: though a completed action is one of an agent's standing answers to the question of what to do, it is not etched in stone, because it forms the basis of a new, related, question, which is "should I have done *that*?" In other words, an action is not an end to thinking but only a waypoint that divides one's continuous thinking about a single question – what's worth doing – into two angles of attack, one open-ended, forward looking, and "legislative" and the other fixed on a specific factual circumstance, backward looking, and "judicial." An action is not itself thinking, but it nonetheless advances it, by setting up the resumption of one's prospective thinking about what to do in a retrospective form.

[38] U.S. CONST. art. 1, § 9, cl. 3.

My claim here is not that we generally act *in order to* set up or spur on our thinking. We act to get by – to get to work and see friends. Rather, my point is that our actions are not automatically definitive of who we are, or constitutive of our life story, since they are the object of further thinking that can alter their role in our identity and their significance in that story.[39]

This kind of retrospective thinking is important to a person's life. Children are often told "think about what you've done." As we know, this is an idiom. It is not a directive to bring to consciousness the mere fact of having done something, as a form of meditation or a way to make oneself sleepy, but is instead, a request to make a particular action the object of open-ended reflection about its wisdom and its meaning to oneself and others. This kind of thinking is a significant part of child development because it is a significant part of an adult life. Being able to reflect on one's actions is an important part of being a responsible member of a moral community and of forming an authentic moral personhood.

The very first step in reflecting on what one has done is to remember doing it.[40] In fact, the remembering is, itself, a bit of self-reflection. It is, in combination with forgetting, an attempt to narrow the field of our cognitive attention in a productive way, so as to focus on what matters.[41]

Forgetting clears up mental space. We automatically forget many banal actions we take, and rightly so. It would be overwhelming and inhibitory to remember them all. However, forgetting a colleague's name or friend's birthday is a *faux pas* because absent a special story, it signals that we do

---

[39] Meir Dan-Cohen, *Revising the Past: On the Metaphysics of Repentance, Forgiveness, and Pardon, in* FORGIVENESS, MERCY, AND CLEMENCY (Austin Sarat and Nasser Hussain eds., 2007).

[40] I focus on memory here, even though, in principle, one may commit a crime in a fugue state or under the influence of psychotropic drugs, not remember it, but then see that one has done it by finding one's video phone was on during the exchange. In this case, one has a justified belief that one has committed a crime but not by remembering that one committed it. I think this belief can be the basis of self-reflection and so does not challenge the arguments made below, though they are all made with regard to a belief about what one has done that is memorial.

[41] *See* O. Carter Snead, *Memory and Punishment*, 64 VAND. L. REV. 1195. 1227 (2011) (discussing the selectivity of memory).

not consider our relationship with them noteworthy. Even significant things may be forgotten in time as we change. This can be for different reasons. We may forget something because we have fully or sufficiently processed and incorporated it into our patterns of action, ways of thinking, or attitudes. This is forgetting when there is not much to be gained from remembering. Alternatively, we may forget because remembering is unproductive, in the sense that it's too painful or depressing. In such cases, we move on, taking whatever lessons we can from what we remember of a particular incident.

Our biographer may be able to remind us later that we did do that horrible/wonderful thing, but our forgetting it is a step toward releasing it from our identity. Consequently, being reminded of something that is no longer consonant with one's identity can be threatening and wrongful, such as being reminded of something embarrassing that one did as a child. For example, setting aside its legality, it was wounding for *The New Yorker* to recount and revisit William Sidis' activities as a child prodigy, because he wished to forget that part of himself, and for it to be forgotten.[42]

The flip side to forgetting what may be released from the mind is remembering what still should occupy us – worth taking with us as we grow and develop. This can be for a number of reasons, encompassing the need to use a password to access our account, but ranging to those actions that we wish to retain for their beauty or emotional power, to those we are ashamed of and so choose to never forget as a warning to ourselves, to those in which we are simply intrigued or puzzled, having done something that we suspect is curious, but may need to be revealed in the fullness of time. It's also important that remembering something is not all or nothing. We may remember things differently at different times, especially if the events we are remembering are

[42] Sidis v. FR Pub. Corporation, 113 F.2d 806 (1940) (child genius unsuccessfully sues publisher as an adult for invasion of privacy after it runs a "where are they now?" article about him).

emotional or especially significant. In changing our feelings about something in the past, we may bring some details into greater clarity or obscurity.

The right theoretical label for the relationship that a memory instantiates and preserves is that of holding responsible. That is, by remembering something, we acknowledge our role in doing it and its concomitant role in changing us. We thereby hold ourselves responsible for it. This is a personal, solitary kind of holding responsible, since it operates on our own continued activity. We carry "baggage" when the thoughts are bad, and we keep someone's memory alive by fondly remembering them. Again, what we hold ourselves responsible for is a proposal for who we wish to be. Others may rightly contradict it in various ways – they may remember us differently – but the existence of such disagreement is significant.

The application of these theoretical points to remembering wrongdoing is likely plain: it is a success to remember actions that seriously hurt others or damage their interests. Someone who attacks another with brass knuckles and then forgets that they did it, or who the victim was, would be monstrously callous. Likewise, it would be especially wrongful to prepare to commit a crime by first taking a memory blocking drug, so that one is not haunted by the memory of doing it, or so that one cannot be a source of evidence about it.[43] Correlatively then, it is a proper and rightful use of one's memory to remember the wrongdoing that one has done, because doing so at least recognizes that this kind of action is not banal or inconsequential, even if one mistakenly thinks that one was justified in doing it. Remembering wrongful action also holds out hope for growth since it preserves the action in one's mind *as* something significant, and so as an eligible target for further reflection.

[43] *See* Snead, *supra* note 41, at 1252 (discussing a similar scenario).

I have argued that thinking adopts a relationship to content that prefigures, in a solitary way, the relationship one will have to it socially and interpersonally, by speaking it or acting on it. Thinking about what one has done is no different. By remembering a past action, one holds oneself responsible for having done it (by recording its occurrence in thought), which is a way of assessing whether to take responsibility for it, outwardly, by asking others to record it too. One may remember doing something and keep it to oneself, but in time, see it as worth revealing to some others. Memoirists are self-consciously engaged in this process, as they not only canvass their memories of what they have done, but then try to decide which of these memories they wish to weave into a public representation of who they are. "To write about one's life is to work out one's identity, never encased in amber."[44]

Here again we see the difference between actions and thoughts. Thoughts that we stop entertaining fall away from us, and others act wrongly if they insist on keeping them attached to us, by holding us responsible for them, but our actions do not belong to us in the same way. They affect a shared world and so are liable to interpretation and preservation by others. Others can be right in refusing to forget what we have done, or of reminding us of it from time to time. Still, how a person thinks about the significance of what she has done is relevant to how others should approach its significance. Apology is a prime example of this, and so renouncing what one has done, can give others warrant in forgiving and forgetting, where forgetting is part of how forgiveness is fully achieved.[45]

I have been focusing on memory as being a bit of self-reflection, but it's also the gateway to more. One does not just remember what one did, but reflects on it and thereby forms new thoughts about it. An important set of such thoughts are the self-conscious attitudes, which are ways of

[44] THOMAS LARSON, THE MEMOIR AND THE MEMOIRIST loc. 1357 (2007).
[45] *See* Dan-Cohen, *supra* note 39.

reacting to what we have done that are explicitly concerned with simulating or anticipating how others will react. For instance, shame is a way of seeing ourselves critically, as we would be seen by others, and spurs us to conceal what we have done and not to take responsibility for it. When we do something embarrassing (a softer kind of shame) we want to be invisible, as we are being held responsible for something that we do not wish to take responsibility for. Guilt on the other hand is the feeling that we have done something wrong such that it must come to the attention of others, for punishment, amends, or other kinds of re-admittance to social good standing.

Guilt is of particular importance, because it represents our conscience, which is the seat of our capacity for responsibility. Conscience guides a person's ability to make choices that she takes to be rightful, and, what is sometimes overlooked, to revise our judgments in light of our contact with a community of others. Guilt is also challenging, because if one feels guilt, one feels the need to make one's actions known to others, and hence to hold oneself open for interlocution regarding the action's merits, while at the same time knowing or suspecting that the action really has no merits – that it is wrong. Hence, when one knows that something is wrong, one takes responsibility for the action only to announce that one is abandoning it as wrongful and so to earn the cooperation of others in excising it from one's identity or record.[46] When one does not know something is wrong, one may first proudly take responsibility for it, only to see it, in time, as cowardly or self-serving.

The phrase "in time" is key because reversing one's attitude toward an action is hard. If one has merely had a thought, it is comparatively easy to revise, because one has not committed oneself to it. It was just a thought. Actions are different because they are worldly memorials to one's one-time judgment about what was worth doing. This point finds support from psychology, which warns us that we tend to treat ourselves as the hero of our own life story, and so to rationalize and

[46] *Id.*

celebrate what we do. The point is not just psychological though. As I said, our actions are termination points for a bit of thinking and we may be held responsible for them, and so we have reason to put significant thought into them and to resist changing our view of them on a dime.[47]

Time is one solution to deliberative inertia. In time, we get perspective, for two reasons. First, we move further away from what we did, which may loosen the grip that our reasons for doing it had on us. We may be able to assess them more critically. Second, we get more information about where our decisions have taken us and how others have reacted. This is a matter of gaining perspective by placing actions within a richer landscape of other decisions. By moving away from actions in time, they stop dominating our vision, and other actions and feelings take up comparable positions of importance too.

The lessons I wish to take from the preceding discussion are nicely summarized by Gilbert Meilaender. He writes, "Each of us is constantly active in memory, constructing and reconstructing the story of his life. We forget some things, of course, as we must. And over time we give new and different significance to events we might once have thought fixed in their meaning; they take on a new shape as the overall shape of life changes."[48]

The specific relevance of these points is that the relationship of a doer to his deed is not fully settled and hence neither is the relationship between a wrongdoer and his wrongdoing. Moreover, it will only be fully settled in the fullness of time, and by the wrongdoer, as he may recognize his wrongdoing on his deathbed or else cling stubbornly to it.[49]

---

[47] Seana Valentine Shiffrin, *Caution About Character Ideals and Capital Punishment: A Reply to Sorell*, 21 CRIM. JUST. ETHICS 35 (2002) (suggesting that ideal agents may take time to come to terms with wrongdoing because their actions will be an expression of considerable effort to think correctly).

[48] Gilbert Meilaender, *Why Remember?*, FIRST THINGS, Aug./Sept. 2003.

[49] *See* Charlotte Alter, *Former Auschwitz Guard Apologizes at Trial: 'I Am Ashamed'*, TIME, Apr. 29, 2016, https://time.com/4312199/auschwitz-nazi-guard-apology-reinhold-hanning/ (94 year old guard confesses to war crimes). I think this example is useful, but we must approach it with caution, as so many of its features are extreme. For

None of this entails that persons cannot be held responsible for what they do, but rather, reminds us that that they must be held responsible for *exactly* what they have *done* – no more and no less – as the meaning of the wrong, for their life and character, is not yet settled at the moment it takes place, and its meaning will depend on what they do about it. As I will argue next, a certain kind of holding responsible, one that goes by way of enlisting a person's very thoughts in the process, holds them responsible for their thinking too. And sweeping up their thinking into holding them responsible is problematic because it unbalances the thinking embodied in a criminal trial and in the criminal defendant himself.

2. Self-Incrimination Corrupts Self-Reflection

In this section I defend the minor premise of my *prima facie* argument, which is that using the contents of a person's memories to convict him of a crime interferes with their proper use, which is as the subject matter of self-reflection.

Gerstein endorses a similar minor premise, but as I mentioned before, his defense of it is sparse. He states that the truth of this premise is the "central difficulty" for his approach, since it requires showing how "protection of the individual's capacity to make *moral* judgments of himself underlies the Fifth Amendment's protection against the compulsion to take part in establishing his own *legal* guilt."[50] This is a prescient formulation of the problem, but the solution offered is strikingly short. The essence of his solution seems to be that "a criminal accusation is necessarily directed to putting [the defendant's] moral guilt in question in his own eyes and in the eyes of the rest of the community" and that "to compel a public admission of what is understood to be guilt by community standards *thus* degrades the individual from his status in the community as a free moral

instance, we may doubt the sincerity of someone who waits several decades to confess. Still such a confession need not be insincere. If it is sincere, my claim is that it changes, to some degree, our view of the action and its relationship to this guard.

[50] Gerstein, *Demise*, *supra* note 13, at 345.

agent; he would have good reason to regard this accusation as a serious affront to his personal dignity and his efforts to come to terms with his own conscience" (emphasis mine).[51] It seems that something significant is missing that would shed more light on the *thus*.

My way of filling in the gap is patterned on the arguments of Chapter 1. The key idea that I need to revive is as follows. Having a thought is way of mentally practicing for a dialogical situation in which one is held responsible for its content. It's like tossing a ball to yourself to get ready to play catch with someone else. Taking responsibility is a way of signaling one's readiness to be held responsible for the thought content in a dialogical situation, as its proponent or spokesperson. This is like throwing the ball to someone else once you feel you have the hang of it; whoever catches it knows they can throw it back to you. Mind reading on the other hand initiates holding responsible and hence the corresponding dialogical situation without such a signaling of readiness on the thinker's part: it forces one into the role of spokesperson for the thought. This is like someone else seeing you tossing the ball to yourself and catching it out of the air, and throwing it back to you, thereby pulling you into a game of catch, despite your unreadiness. The game won't go well, nor will any dialogical situation involving a thought for which one has not take responsibility.

An action is different than a thought. Hence, one may be held responsible for an action without any initiation of a dialogical situation between two thinkers. DNA analysis, witnesses, or a security camera may reveal what one did. One may be, in these ways, held responsible for what one has done before one takes responsibility for it, without any kind of damaging prematurity or interference with thought. The reason is that the action is the end point of one's preceding thinking,

[51] Gerstein, *Demise*, *supra* note 13, at 353. Another minor puzzle is why he says that the *accusation* is an affront to come to terms with his own conscience. Is the claim that even accusing someone will interfere with one's status as a "free moral agent?" In place of "accusation," I think the noun one expects is "public admission."

and so establishing it with evidence that does not include the defendant's thoughts does not endanger any thinking that is in progress or still to be done.

However, an attempt to find out what someone did *by* using her memory of doing it *does* create a dialogical situation, because the vehicle for gaining the information is a thought. In such cases, the pattern outlined just above recurs, not for the content of the thought itself, but for the action that the content represents or describes. That is, by remembering what one did, one is mentally practicing for a dialogical situation in which one is held responsible for the doing of it. Taking responsibility for what one did is a way of signaling one's readiness to be held responsible for the action in a dialogical situation, as its proponent or spokesperson. This is accomplished by confessing, which is sincere speech consisting of the content of the relevant memory, offered *for the purpose of being held responsible for the action*.[52] Now the key point and parallel: to read someone's memory of committing a crime *for the purpose of holding them responsible for committing it*,[53] creates the same dialogical situation – it marshals the same mental content in the same way as a genuine confession – but without a signaling of readiness to be in that situation by the guilty party.

The result is this. When a defendant's memories are used, via mindreading or through compelled speech, to hold him responsible for the action he remembers, a dialogical situation is created. This situation is the same as that which would be created if he confessed, i.e. one in which

[52] Some such sincere speech is better termed an admission rather than a confession, when it is not offered for the purpose of being held responsible for the action. For instance, a terrorist organization may issue a press release "taking responsibility" for an attack, but this is a (proud) admission, designed to strike fear in others or garner new recruits. It is not a confession because it is not made for the purpose of being held responsible by the legal system of the victims or the international legal system.

[53] This is because one may read someone's memory, not to hold them responsible for the action it represents, but just to find out what they did. I am not claiming that this would be innocent, but only that it is distinct. For instance, a scientist may read someone's memory to find out what they did on Tuesday, without any interest in the rightfulness of what the memory reveals that they did. This is the idea of immunity in a nutshell, though it can find expression in confidentiality or immunity to discipline, etc. For example, a lawyer may say to a client "tell me everything, you won't be in trouble" so as to get the facts needed to defend the case. A college administrator may say to some freshmen "tell me about your parties, you won't be in trouble," in order to get information about underage drinking on campus.

he is a spokesperson for his action, and can be asked to answer the question "what do you have to say for yourself?" However, when his memories are accessed involuntarily, e.g., by a prosecutor's mind-reading machine, then he has not signaled a readiness to be such a spokesperson and so to *have something to say for himself*. He may in fact be ready to be a spokesperson for his action, but given the consequences of being unready (to be outlined next), proceeding without such a signal is wrongful.

The consequence is a foregone opportunity to consider the relevant moral subject matter in an unmediated fashion. Put another way: being denied a chance to settle for oneself one's emotions and judgments regarding what one has one, before being asked to inject that view into a dialogue with others risks creating a gap between oneself as a thinker and the subject matter that one has an interest in thinking about, where the gap is occupied by the thinking of others. One is made party to an interpersonal deliberation about one's actions rather than being able to come to that discussion with one's considered position. The latter may of course be refuted or challenged, but one ought to have a position first, for such refutations and challenges to be sites for rational response and attention rather than mere psychological forces or browbeating.

Let me give a name to the dialogical situation that arises when one's way of holding oneself responsible for a crime – one's memory of doing it – is known and used to hold one responsible. It is a *confessional situation*. Note that confessional situations encompass both those in which one voluntarily confesses and those in which one confesses under some kind of pressure to do so, such as police interrogation tactics, or punishment for not telling the truth. They also encompass situations in which one's memory is accessed by oneself and made available through speech, and those in which one's memory is accessed by mind reading technology or telepathy, etc. This latter

point is important as most actual confessional situations rely on speech given the lack of highly precise mind reading capabilities, but the kind of wrong I'm trying to isolate goes beyond them.[54]

Take confession as a (Catholic) religious practice. "The act of confession must be considered private, protected, sacramental because it is so potent. It carries a potential use value which, in the Christian religious tradition, is carefully policed."[55] The potency owes to the fact that someone who works through their wrongdoing with another enacts a shared understanding of what they have done and its significance. The priest mediates the process of repentance and making amends, offering a lens to see that journey, vocabulary for how to think about it, and guidance on how to accomplish it. He instructs one in examining oneself and so, in making oneself, especially in the decision about what sorts of prayers (cognitive items) to use to aid in penance.[56] This is a powerful position to occupy, since it consists in being a cognitive fellow-traveler (and authority) in a very fundamental aspect of moral life. Owing to the power, it is private, in the sense that it would be risky to engage in it with just anyone. A priest is selected by a Catholic as having the right character and training to receive a confession.[57]

Is confession to a priest voluntary or not? It may have elements of both. One may go to a priest when one is ready to confess, where this readiness entails a fully fleshed out view of what one did and how to remedy it. However, in Catholicism, the spiritual experts are priests, and so a

---

[54] Perhaps God knows what one has done and knows one's memory of doing it, and so one is, in this way, always in a confessional posture toward God. There are complexities here better handled by religious scholars. Margaret Falls-Corbitt and F. Michael McLain, *God and Privacy*, 9 FAITH AND PHIL. 369 (1992); Susan Kramer, *"We Speak to God With Our Thoughts": Abelard and the Implications of Private Communication With God*, 69 CHURCH HISTORY 18 (2000).

[55] PETER BROOKS, TROUBLING CONFESSIONS 89 (2000).

[56] *Id.* at 94 (discussing the historical evolution of priest as shepherd who "keeps tabs on one's spiritual condition, exercising both reprobation and healing."); CATECHISM OF THE CATHOLIC CHURCH ¶1455 (1997) (available at https://www.usccb.org/sites/default/files/flipbooks/catechism/VI/)("The confession (or disclosure) of sins, even from a simply human point of view, frees us and facilitates our reconciliation with others. Through such an admission man looks squarely at the sins he is guilty of, takes responsibility for them, and thereby opens himself again to God and to the communion of the Church in order to make a new future possible.").

[57] Gerstein, *Privacy*, *supra* note 13, at 92 (privacy of confession in history and importance of the identity of the confessor); Christine P. Bartholomew, *Exorcising the Clergy Privilege*, 103 VA. L. REV. 1015, 1060 (2017) (surveying uses of penitent's privilege and arguing it is being less recognized by courts).

decision to confess may also reflect a residual uncertainty about how to address a sin. One may know that one has sinned and that this requires confession, as per doctrine, but one may wish to talk with the priest to understand further – to hear in more detail what the proper response should be, even if one knows roughly what the Christian remediation should be from sermons or the Bible.[58] The point is that the state of having confessed, to the degree that one does so seeking instruction, is a position of self-plasticity, and one that legitimates giving great control to persons over whom they confess to.

If we turn to more troubling religious confessions, such as those produced by the inquisition, we see the same logic, which is that forcing someone to submit their own thoughts as the basis for punishment is a way to make people more malleable and compliant. In a confession to the inquisition, one is not voluntarily offering the contents of one's guilty mind to receive instruction from a spiritual advisor. Rather, one is forced to bring one's guilty mind out in the open, to receive punishment. Crucially though, in confessing under torture or prolonged interrogation, despite being starkly involuntary, a sinner is nonetheless thought to start down the road of penance, and this is because admitting guilt to a moral authority initiates a confessional situation, which the inquisitor can exploit.

In this example, we see how confession "offers a royal way into the individual's psyche, into his or her personal beliefs…but in the mode of social and ideological pressure, as the policing of individual belief."[59] When one confesses voluntarily, to a trusted advisor, one may be educated. One's moral selfhood is receptive to being built up and consolidated. When one confesses involuntarily, one's moral selfhood is similarly ready to be rebuilt, but in a perverted or mock version

[58] BROOKS, *supra* note 55, at 94 ("Confession and preaching indeed move forward in tandem: the individual's self-examination and avowal of sin is accompanied by exhortation to live according to Christian example.").

[59] *Id.* at 99.

of the voluntary case. One is not ready to be educated, but rather to be reeducated in the totalitarian sense of that word.

Other examples illustrate the plasticity of moral personhood under confessional stress.[60] Consider psychoanalytical practice. The psychoanalyst is not interested in the "guilty" actions that the subject readily confesses. Rather, he seeks out places where the subject is resistant and guarded, attempting to gain what can be understood as "coerced" or "involuntary" confession by way of guileful and probing interlocution (not unlike a cross examination).[61] He seeks to catch the subject out, and this catching out reveals the subject's mind in a way that she is unready to rationalize, defend, or understand, and thereby creates an opening for positive change to occur – for a psychological shift.[62]

Again though, this state of openness carries risks, as it makes the patient suggestable or vulnerable to the personality and authority of another. Consequently, the psychoanalytical patient is entitled to privacy and confidentiality. Trust in one's specific therapist is also necessary, but at the same time poses its own problems given the nature of the patient's vulnerability. Much psychoanalytic practice is aimed at ensuring that openings in the mental life of patients are not exploited by the analyst, which can result in a dark version of trust that could be described as a form of dependency or a faux-intimate relationship.[63] The analyst relies, "like both priest and police

[60] The examples also tend to show the more general truth that dialogical situations (or deliberative situations with others more broadly) and their features powerfully shape the ways thinking proceeds within them. In this way, the arguments here can be understood as bolstering some of the claims in chapter 1.

[61] BROOKS, *supra* note 55, at 117.

[62] *Id.* at 116 (quoting Freud: "In Confession the sinner tells what he knows; in analysis the neurotic has to tell more.").

[63] For instance, eminently sensible rules about forming romantic or sexual relationships with patients. *See* AMERICAN PSYCHIATRIC ASSOCIATION, PRINCIPLES OF MEDICAL ETHICS, WITH ANNOTATIONS ESPECIALLY APPLICABLE TO PSYCHIATRY 2.1 (2013) (available at https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/Ethics/principles-medical-ethics.pdf) ("Sexual activity with a current or former patient is unethical.").

interrogator, on a certain transferential bond with the analysand" but must employ that bond skillfully and carefully, lest it cause fresh psychic damage.[64]

Religious confession and psychoanalysis demonstrate my main point, which is that forced confessional situations are a harmful parody of voluntary confessions, though both operate by "opening" the confessant to the thinking of others. The latter may lead to moral growth, as one's authentic and robust, though possibly incomplete, self-reflection can be seasoned with the wisdom of trusted others. By contrast, forced confessional situations interrupt one's authentic and robust self-examination and so replace what could have been real dialogue with faux-dialogue that can be better described as the transference of ideology from the confessor to the confessant.

To apply my discussion of confessional situations to self-incrimination, notice that the power of the former depends on how "real" or vivid it makes a dialogical situation and how immersive it is for the confessant. As we have seen, one important feature of making a confessional situation vivid is to have some actual dialogue, and so require one party to say something to another. Speech activates one's mind and so may produce pressure to adopt the thing signified (a thought) by compelling the creation of its sign (speech).[65] For instance, Saint Bonaventure thought confession had to be spoken because the shame is greater that way.[66] Other features involve the application of coercion or its threat, the presence of authority figures, control over the direction of the dialogue, and the audience.

All of these things, it should not be missed, describe a trial. In a trial setting, one is relentlessly questioned – made to answer under pain of punishment – by a public official in a public

[64] BROOKS, *supra* note 55, at 116.

[65] Seana Valentine Shiffrin, *What is Really Wrong with Compelled Association?*, 99 NW. U. L. REV. 2 (2005) (arguing that compelled speech risks distorting a person's thinking); West Virginia Bd. of Ed. v. Barnette, 319 U.S. 624 (1943).

[66] BROOKS, *supra* note 55, at 95.

forum, on matters of great concern to the community.[67] These features of trials are justified, but they create a vivid confessional situation for defendants and hence risks sabotaging their independent self-examination regarding their crimes.

A criminal trial can be thought of as a solemnized and ritualized stretch of public thinking, in which a record of what happened is collectively assembled by witnesses in the presence of the jury, and evaluated with reference to the law. Consequently, forcing the defendant to testify forces him to participate in a vivid, collective confessional situation for which he may not be ready. He would be forced to participate, along with society as represented by the prosecutor, in creating a collective account of what he has done, its significance, and its proper remediation. These judgments cannot help but stifle the defendant's independent answers to those questions. A trial is, by design, an activity of meaning-making at scale, and so, without the benefit of a space in which to stand aside, apart, and away from the enactment of one's guilt, through the reasoning and collective thinking that constitutes a criminal proceeding, one's own feelings and reactions will be swallowed up or overborne, and replaced by the judgments of the community.

Further proof of this fact is that show trials are attempts to do to a populace what forced confession does to a particular defendant, by harnessing the latter. By making the public a party to a confessional event in which the defendant takes responsibility for invented crimes, the state attempts to make its citizens witness the total capitulation of one of its enemies, and thereby to warn against dissent, but more importantly, to legitimate itself. Compare a show trial with something of its opposite – a trial for a civil disobedient. Here, a citizen has acted against the law "in the open,"

[67] Some scholars argue that testifying as a defendant in a criminal case is such a high-stakes experience, that it may intimidating for someone who is *innocent*. *See* Schulhofer, *supra* note 4 (arguing that the privilege against self-incrimination is for protecting the innocent from making discrediting missteps on the witness stand).

confessed, and "accepted his punishment."[68] This is a complete and voluntary taking of responsibility, and its effect, rather than to allow ideology to be poured into the individual or to legitimate the regime, is to confront the regime with the individual's undiminished and unyielding sense of justice, and correspondingly delegitimate the regime as failing to live up to the ideal of the rule of law.

If we accept that there is a serious privacy interest in standing aside and apart from a trial, understood as collectively enacted thinking, then we may also add nuance to our understanding of other relatives of the privilege. Take, briefly, the right to have an attorney in a criminal trial and the necessity of her presence when being interrogated by the police. It is important for many reasons, not least because it lends the attorney's tactical and legal expertise to a defendant who is facing the assembled legal power of the state. This expertise gives the defendant a fighting chance and so puts the government to its proof.

However, we can also see the attorney as partly fulfilling a privacy role. She provides, as a corollary to the defendant's right not to talk with a state, a guaranteed person to whom he *can* talk. After all, the defendant can confess to his lawyer, in the way that a sinner may confess to a priest, as a way to tap the wisdom of a representative of the legal system before interacting with it more broadly. The defendant needs some private space for his conscience to stretch out and make contact with another mind, as part of guaranteeing the full value of having a right to close himself off from the government's proceedings against him. Priests and spouses may help in this regard, but not everyone has these relationships. A lawyer is something that may be guaranteed to each person and serve as a secure partner to think with.

[68] *See* Bruce Ledewitz, *Civil Disobedience, Injunctions, and the First Amendment*, 19 HOFSTRA L. REV. 67, 71 (1990) (civil disobedients accept punishment).

My way of summarizing this section is by considering an objection that brings out its core claim. The objection asks whether the appeal to the psychological effects of a criminal trial proves too much: why wouldn't the ordeal of the trial itself risk distorting the defendant's self-reflection? Perhaps the defendant does not have to testify during the trial, but he is there, and he has to witness the machinery of justice deciding his legal culpability. If witnessing this could interfere with self-reflection, then doesn't it follow that trials themselves risk damaging a serious moral interest of the defendant?

My response to this objection is first, to agree that there are significant psychological effects of being put on trial. Second, to point out that a defendant may choose to avoid them by choice. But third and most importantly, to contend that being present at one's own trial is importantly different from *participating* in it as a confessant/witness.

Defendants do not have to be at their own trial.[69] They may be tried *in absentia* if they knowingly and voluntarily waive their right to be present.[70] So, there is a serious difference between being required to testify and being, as a default, present at one's own trial. Still, this point only superficially addresses the objection because defendants may be under pressure to attend, due to the risk of prejudicing the jury if they absent themselves, or the discomfort of staying in a cell during the proceedings. Furthermore, a defendant may not be self-conscious enough to consider the question of whether to attend his trial and so attend because that is the default. Nevertheless, it is significant that a defendant's presence at trail is not compelled.

---

[69] Eugene L. Shapiro, *Examining an Underdeveloped Constitutional Standard: Trial in Absentia and the Relinquishment of A Criminal Defendant's Right to Be Present*, 96 MARQ. L. REV. 591, 618 (2012).

[70] Smith v. Mann, 173 F.3d 73, 76 (2d Cir. 1999) ("we hold that nothing in the Constitution prohibits a trial from being commenced in the defendant's absence so long as the defendant knowingly and voluntarily waives his right to be present.").

A deeper response to the objection draws on the reasons why it is often seen as good, though not obligatory, for a defendant to attend his trial. Those reasons suppose the defendant to be playing a role that is in the opposite of the one he plays in a confessional situation. If this is true, then attendance itself does risk distorting the defendant's thinking in the same way. A privacy violation, including a confessional situation, is one in which one is held responsible for a thought before one is ready. One is the object of responsibility practices. However, one's role as a defendant who is present in his own trial is as the subject of responsibility practices. That is, the defendant's role in his own trial is as an auditor. He is there to enforce fair procedure, to observe, and to be heard on his own terms, through his lawyer.[71] His right to silence is thus not only to keep him from being forced to enact his own guilt – to be a subordinate participant – but also, to affirm his role as an adversary, which is to say, as an equal participant in the trial, on the same footing as the government.

For these reasons then, I think that the character of a defendant's trial participation can strengthen the independent thinking of the defendant and to give it some measure of affirmation. In fact, the right to silence is a key part of making the nature of the defendant's trial participation salient. His right to silence is a part of his entitlement to be an observer and, to some degree, sit in judgment of the proceedings.

Having outlined the confessional risk at the heart of self-incrimination, I turn to analyzing how the law should shape itself in response to it.

C. Applying the Argument to Legal Doctrine

In the last section, I argued that confessional situations risk interrupting and compromising the self-reflection of those who are made to confess. Because testimony at one's own criminal trial

[71] *See* Shapiro, *supra* note 69, at 619-620.

enacts a confessional situation, it risks interfering with the defendant's self-reflection, and this risk weighs in favor of not requiring such testimony. Preserving the opportunity for individual moral self-reflection is an important justification for the privilege against self-incrimination. Still, it's just one moral consideration in a universe of competing moral and legal imperatives. Consequently, more needs to be said about how the danger of distorting moral self-reflection should be accommodated by legal doctrine.

The most urgent reaction to address is that the principle I have identified simply does not deserve to be weighed very heavily, if it all, in thinking about how to respond to those who commit serious wrongs. The reaction can be put bluntly by asking why the readiness of the defendant should be taken into account. After all, the defendant has acted in contravention of publicly arrived at rules about reasonable behavior, and has done so at his own convenience. The public must now pick up the pieces of his destructive choice, so why cater to his reflective needs? Moreover, if the defendant could be forced to testify against himself, this may spare the plaintiff from having to testify about the wrong she suffered (and so perhaps avoid interfering with her own coming to terms with what she has suffered), and may provoke more robust reflection from him to boot.

A multi-part response is needed. First, I wish to resist the view that the defendant's opportunity to reckon with their wrongdoing is of little significance in itself. Persons who have not committed a crime have a serious interest in being able to reflect about what they have done and how to interact with the corresponding memories on their own terms. Doing so is a part of creating a biography that has oneself rather than others at the core of it. It's an interest of a foundational kind – an interest that is nothing less than that of making contact with the world and being in it through the activity of, and so as, oneself. This interest does not disappear when one commits a criminally recognized wrong. No doubt, the wrong warrants the community in responding, but the

nature of this, of having a warrant to respond, does not supplant or diminish the interest that persons have in their thinking, but presumes it. After all, if there is a justification for prosecuting persons, then persons must be able to accept it as justified, by thinking about it along with what they have done. If the opportunity for this thinking is compromised, then so is the presumed addressee of justified criminal punishment.

A different way of putting this point appeals to the nature of holding responsible. As I have argued at different points, holding persons responsible for something is a way of showing care for their thinking, and so, partly, sharing responsibility with them for making it good. Indeed, I think that states have an obligation to improve the thinking of their citizens, but only by providing an environment in which it will thrive. One way the state discharges this obligation is by holding persons responsible for thinking that takes the form of illegal action. That is, once a person acts illegally, they have shown that their thinking is in need of help, and this in turn triggers the state's obligation to hold them responsible, as a way of thinking with them.

Crucially, thinking with someone is aimed at improving her thinking and not supplanting it or replacing it. "To act humanely requires … a special attitude on the part of the judge, jailor, or therapist. A person must be thought to be in rightful possession of his desires, needs, and beliefs, however much we may wish him to change them or give them up, or however deeply we feel them wrong or bad."[72] A second way of putting this point is that "If I control somebody, there is no point of making that person accountable to me. Accountable for what? For the things I induce him or her to do? … Accountability concerns agents, not subjects…Or, more precisely, it concerns subjects

[72] A.R. Louch, *Scientific Discovery and Legal Change*, 49 THE MONIST 485, 499-500 (1965).

only as far as we ascribe some degree of freedom to them."[73] Persons are held responsible by the law, which presumes them to be thinkers in their own right.

The point of the foregoing is not that the interest a defendant has in coming to terms with his wrongdoing is absolutely supreme, but only that it remains significant, despite his wrongdoing. It is an interest that the law presumes when it holds defendants responsible for their wrongdoing, rather than responding in some other way. Being held responsible is appropriate for someone who can sit in judgment of themselves, through the exercise of a conscience, and this capacity is exercised over time. The failure of it at one moment requires a response, but a one-time failure is not a failure of the capacity or the forfeiture of the interest in exercising it.

Moreover, it turns out, the defendant's interest in self-reflection is not forced to stand alone in supporting a categorical prohibition on compelling defendants to testify. Instead, the privilege acts like something of a mortar, by linking and shoring up other legal protections for thinking. Viewed in isolation, the shoring up functions may seem unimpressive, but the whole of them is greater than the sum of its parts.

For instance, the privilege against self-incrimination helps bolster the thought crime doctrine, conceptually and in practice. Conceptually, the privilege against self-incrimination helps to draw attention to thought crimes that do not announce themselves by flatly omitting an *actus reus*. Consider a law that requires persons to report any criminal intentions that they are harboring (like an intention to rob a bank). Such a law does not criminalize the criminal intentions themselves. It has an *actus reus*: failing to report one's criminal intention. Nonetheless, it is plausibly a thought crime because it criminalizes the failure to do something that violates mental privacy; namely, revealing

[73] Andreas Schedler, *Conceptualizing Accountability, in* THE SELF RESTRAINING STATE: POWER AND ACCOUNTABILITY IN NEW DEMOCRACIES 13, 20 (Andreas Schedler, Larry Diamond, and Marc Plattner eds., 1999).

one's thoughts. This may be acknowledged on its own, but the privilege against self-incrimination draws attention to the danger that this law poses to thought. After all, the imagined law would require persons to supply the authorities with evidence of their forthcoming wrongdoing.

The privilege against self-incrimination also reinforces the thought crime doctrine in practice, because wrongful attempts to prosecute persons for their thoughts will require evidence of the same, and such evidence will often reside in the mind of the accused.[74] David Dolinko argues that this kind of reinforcement is of minimal importance since, he says, persons who harbor criminalized thoughts will likely have expressed them to someone, and if they have not, and have kept them completely to themselves, they can lie about what they believe if forced to testify. I'm doubtful of both claims.[75] Persons may keep criminalized thoughts to themselves or communicate them in coded or anonymous fashion to others. Moreover, some thoughts crimes may not target ideologies that are shared among groups of persons, but involve fantasies or intentions that persons would likely hide entirely from others if they were criminalized. As for lying, it's not a costless solution, since a regime may succeed in demeaning and shaming free-thinkers by getting them to lie to cover up their beliefs. Additionally, new technology may make it easier to detect when someone is lying. Hence, there is room for the privilege to play a role in restricting government overreach into the mind.

The justice of the adversarial system more generally is also enhanced by the privilege against self-incrimination, because the latter removes one incentive for using abusive interrogation techniques while also setting the moral tone of the justice system by constraining the judiciary from using contempt to extract a confession. Dolinko also doubts this argument. He contends that

[74] Courts in the 1950s were reluctant to enforce First Amendment protections for Communists and so many resisted inquiries into their beliefs by relying on the privilege against self-incrimination. *See* Dolinko, *supra* note 1, at 1085.

[75] *Id.*

torture is already barred by the due process clause and that persons can be tortured into waiving the privilege. These are possibilities, but there is a lot of ground between scrupulous government obedience to the prohibition against torture and the use of torture at all costs, heedless of any consequences and in circumvention of all legal standards.

Indeed, I think the risk of abusive interrogation tactics probably lies more in the middle ground, where shortcuts are taken, where moderate callousness and laziness are factors, and a lack of institutional concern plays a role in encouraging the former. The privilege may combat these forces by setting a clear and unequivocal right that is likely to be understood by as wide an audience as possible, both defendants and various officers of the justice system. Its endorsement by the judiciary can be recognized by those involved, it can be invoked by defendants, and it may provide some discouragement to the easy path to secure a conviction. A defendant could be tortured into waiving the privilege, but he equally well might reveal what has been happening in open court. To the degree that the value of the right to silence is upheld, defendants may be more confident doing the latter.

A final related point is about symbolism. Charles Fried argues that "By according the privilege as fully as it does, our society affirms the extreme value of the individual's control over information about himself."[76] I agree, but wish to distinguish my argument from his. I think Fried is right that the privilege against self-incrimination is useful as a simple cognitive touchstone for a society. It is partly its simplicity that helps it collect and sum up various different concerns a society might have about interferences with thinking. However, I see this function as being performed in addition to, and because of, the way the privilege vindicates a person's interest in moral reflection as well as reinforcing other legal values. Fried however argues that the privilege may not be justified, and functions, more or less, *only* as a symbol. I think his way of conceiving of symbolism undersells it,

[76] Charles Fried, *Privacy [A Moral Analysis], in* PHILOSOPHICAL DIMENSIONS OF PRIVACY 203, 214 (Ferdinand Schoeman eds., 1984).

because symbolizing a value with a practice that is not justified by that value invites cynicism. Isn't the symbol just a monument to a mistakenly excessive deference to the value? Why not use a different symbol; one that is less costly? My way of avoiding these objections is to insist that the symbolic role of the privilege is a matter of packaging up, for easy consumption and appreciation by a society, a range of sufficiently justificatory benefits to having it, that are linked by their attempt to respect the thinking of those who are otherwise properly subject to criminal law.

I have argued that the privilege against self-incrimination serves the important interest of respecting the self-reflection of the defendant, but not just that. The privilege also practically reinforces and makes doctrinally salient other legal protections for thinking such as the thought crime doctrine and the prohibition against torture. It thus contributes to an overall legal regime in which the urgency to punish persons for serious wrongs does not cross into totalitarian territory.

I will now shift gears to a different set of objections to the mental privacy approach. These objections argue that honoring the importance of mental privacy would entail that the privilege should be implausibly extended to cover ordinary fact witnesses, immunized witnesses, and civil defendants.

Ordinary fact witnesses are obligated to reveal the contents of their mind, and this is a sound obligation. But, an objection goes, how can my view accept this obligation, given that fact witnesses have to reveal the contents of their mind before they may wish to? My answer relies on when persons can be expected to be finished with thinking, where the answer varies by the kind of thinking at issue.

For some kinds of thinking, it is not reasonable to require persons to be finished with it. Thoughts about what is true regarding subject matters like politics, religion, math, or art, have no clearly demarcated endpoint, and so there is no ground to say that *now* someone should be

comfortable revealing what they think. Similarly, reflection on one's past actions has no clearly demarcated endpoint, after which one must be ready to take responsibility for them. Such thinking is a matter of settling who one is and so is always changing. Moreover, taking responsibility for wrongdoing is especially difficult, because one must reverse previous thinking. and self-criticism of this kind is deeply uncomfortable. For these reasons, I don't think that the privilege against self-incrimination should be time limited, so that the government could require testimony if the crime was committed, say, ten years in the in the past. Any time limit would be arbitrary.

By contrast, thinking that consists of using the senses to form experiential beliefs and memories is different.[77] Thinking of this kind is automatic and similar across persons. Hence, it is reasonable to expect citizens to be finished with this kind of thinking immediately or soon after the relevant memories or beliefs are formed. This is not to say persons must be ready to take responsibility for the truth of what actually happened to them, but only something weaker, to be sincere in reporting what they experienced as happening, which may be vague or permissibly hedged – "the bus that hit the pedestrian? It might have been blue or red. I couldn't tell it was too dark and I was tired." Since fact witnesses are only obligated to testify to their firsthand knowledge, their obligation to testify is unobjectionable from a mental privacy perspective.

A different objection concerns immunized witnesses. Witnesses who refuse to testify on the grounds that their testimony may be incriminating may nonetheless be forced to testify if they are statutorily granted immunity from prosecution that is commensurate with the protection afforded by the privilege.[78] This practice seems sound, but mental privacy arguably condemns it, since a person will be compelled to admit a serious wrongdoing to the community when they are unready to take

[77] The law recognizes this in the form of the lay witness / expert witness distinction. Ordinarily, an expert may not be forced to give an opinion, because, on my telling, this is asking for a more robust use of her mind then asking a lay witness what they heard and saw.

[78] Kastigar v. United States, 406 U.S. 441 (1972) (discussing permissibility of immunity statutes).

responsibility for it. Put succinctly, immunity practices suppose that self-incrimination cannot be compelled because of the legal consequences that stem from incrimination, whereas mental privacy supposes that it cannot be compelled because of the privacy consequences of self-incrimination itself.

Others have argued that immunity statutes are not permissible and should be reformed.[79] I am sympathetic to such views, but I am unsure they carry the day. One consideration that is relevant is that though the lack of punishment is not itself relevant to mental privacy, the fact that one will be spared it may give one greater control over how one reflects on one's wrongdoing. Of course, one may go back to a life of crime, but with some ability to structure interactions with friends and advisors as a free citizen, robust moral reflection may be possible. This possibility is made more likely by the fact that in testifying about what one did, the purpose is not to hold oneself responsible but another.

Another common objection to the mental privacy rationale for the privilege against self-incrimination is that civil defendants are required to testify about their civil wrongs. So, the thinking goes, there cannot be a privacy rationale at work in the privilege against self-incrimination, otherwise it should apply to civil defendants, on the assumption that their privacy matters just as much as that of criminal defendants.

It may seem that my view provides an easy response. Self-incrimination interferes with one's moral reflection because crimes are immoral, but civil violations are not. Hence, testifying about one's civil actions does not interfere with any moral reflection. This is too easy though, because it

[79] *See* Gerstein, *Demise*, *supra* note 13, at 355-56 (doubting that immunity completely neutralizes the moral risks of compelled self-incrimination); Richard Wasserstrom, *Privacy: Some Arguments and Assumptions, in* PHILOSOPHICAL DIMENSIONS OF PRIVACY 317, 322-23 (Ferdinand Schoeman, ed. 1984) (arguing that immunity statutes are "properly subject to criticism").

rests on the mistaken assumption that civil liability has no relationship to morality. It does though. Civil law concerns obligations to take reasonable care, keep promises, maintain non-discriminatory policies, and more. Consequently, I think that testimony by defendants who have committed a civil wrong may initiate a confessional situation, and so, distort their moral self-reflection.

Nonetheless, I think there is a difference in civil and criminal proceedings that makes a difference to how testifying in them affects one's own moral reflection. The difference I think is fairly uncontroversial, which is that a criminal trial has, as one of its parties, the people as a collective, whereas a civil suit, though it may involve a government officer or entity, does not. The people, represented through the prosecution, are a necessary party to criminal convictions because crimes are offenses against the community as such. An important symptom of this difference is that a civil plaintiff is not required to only consider the interest of the polity as a whole in abandoning or pursuing her suit. For instance, if a doctor makes a mistake and harms a patient, the patient may drop the suit if she is satisfied with the doctor's apology. However, if the victim of a criminal wrong (if there is an identifiable victim) accepts the perpetrator's apology, this may be important, but it does not settle things, because there is still a matter of deciding what the public's interest is with regard to the case and the remedy sought, etc. Punishment is not imposed so that the victim can be made whole, but to pay a debt to society.

This difference makes a difference, such that criminal prosecution poses a distinct threat to the moral self-reflection of the defendant, for two reasons. The first is that kind of moral self-reflection that is at issue with a collective wrong is more daunting and difficult than when there is an identifiable party or parties whom one has wronged. A civil wrong is a wrong that is recognized by the community as wrong and that the community finds it worth its attention to remedy, but the

needed moral repair work is not with the community as such. With a criminal wrong, the community as such is one's victim and the entity who is requiring one to testify.

The second is that there is a greater possibility of confusing law with morality where the legal norms at issue are those that constitute violation of the community at large. Criminal norms enjoy a high degree of agreement and are zealously enforced, thus those who speak on its behalf enjoy an especially powerful degree of moral authority. Part of the power of the inquisition came from its religious affiliation, and hence its claim to speak on behalf of a community of faith and so god. In the face of organized communal norm enforcement, the possibility of confusing community norms with morality itself, and so shortchanging one's direct consideration of the underlying moral reasons against, say, tax evasion, or murder is more likely. Privacy in general works this way. Exposure of one's dissident views is most damaging when they are exposed to the public at large, because such exposure brings to bear the weight of such wide agreement and judgment as to one's own thoughts.

An additional point is a backup argument. Let me assume that there is no distinct risk of distorting self-reflection that is posed by criminal as opposed to civil testimony. Still, it is the criminal laws that society is most likely to enforce too zealously and too totally, as they are, again, the supposed norms that protect society as such. They are likely to enjoy enormous consensus and correspondingly, the criminals are a category of persons that is perennially at risk of demonization or degradation. Defendants awaiting trial are often in the custody of the government and so highly vulnerable to its activities. Moreover, criminal law violations are the most likely to draw the exertions of prosecutors and police interrogators, and mostly likely to motivate the need to hear about the defendant's crime from his own lips. Due to criminal law being a predictable site of government excess, a self-incrimination privilege may be most justified for criminal proceedings.

III. Conclusion

The privacy rationale for the privilege against self-incrimination has been out of fashion for some time, partly owing, I suspect, to the difficulty in understanding why privacy is valuable, especially mental privacy. Nonetheless, I think it is also a very plausible rationale to pursue, because it seems that in principle, the government will be able to access the minds of suspects by observing their brains; that this kind of access will be completely unobjectionable according to many non-privacy rationales for the privilege and yet intuitively circumvents it; and that this circumvention is an invasion of privacy. If this is right, then it seems that the privilege as it has operated in history has been to bar one way of accessing a defendant's incriminating thoughts (through coerced testimony), but that it is this access itself, however achieved, to which the privilege applies.

I have tried to show why limiting such access is justified. On my telling, it is justified by the need, honored by many criminal law doctrines, to ensure that holding persons responsible for doing wrong does not slide into a campaign against their personhood. The *actus reus* requirement is part of this commitment, as is the First Amendment, but these are most prominently applicable to the content of the criminal laws rather than to the way in which they are enforced. The self-incrimination privilege carries through the values of these doctrines to the aftermath of criminal wrongdoing.

By acting wrongly, one contravenes the judgment of the community about how to structure interactions in a common world. Hence, it is the public's business to hold persons responsible for crime, even if, as is often the case, they do not take responsibility for its commission. However, the risk of having this responsibility is overzealousness, such that the state seeks to hold persons responsible so thoroughly that they start to inhibit the ability for them to take responsibility. This is objectionable. Persons have, in principle, in the fullness of time, the ability to find their way back to

the moral community, and they have a preeminent interest in this road of return being kept open, most of all by the state.

Does this mean that the privilege is justified only by the needs of the guilty? A yes answer could be defended by arguing that an innocent defendant is just a fact witness and so could be required to give her answers to various questions. After all, requiring an innocent defendant to answer could prevent a mistaken conviction.

I'm not sure though. Although a defendant may not have committed the crime that he is accused of, he may have acted badly or have been engaged in something embarrassing. If such defendants were not covered by the privilege, it might encourage prosecutions that have the aim of getting someone to speak such truths, though there are a number of checks that would arguably prevent fabricated charges. A second argument is that defendants, even if innocent, are not ordinary fact witnesses, because they are being asked to secure their own innocence. Though they have every prudential reason to do so, there may be unfairness in placing someone's thoughts at the core of their own defense against criminal charges. Such unfairness may be found, in an attenuated form, in some misprision statutes that require persons to report crimes that they witness. If not properly restricted, such statutes objectionably "impress" ordinary citizens into law enforcement officers.[80] These arguments need more exploration, but I think there is room for finding a rationale to the privilege that is enjoyed, in different ways, by the innocent and the guilty. This chapter, however, focused on how it might have a moral justification, even as invoked by the guilty.

[80] Roberts v. United States, 445 U.S. 552, 570 (1980) ("we do not give…[public] officers the authority to impress the citizenry into the prosecutorial enterprise.")

Conclusion

Some violations of privacy consist in minding the business of another, or, in showing concern for something that is properly the concern of another. This dissertation argued that there is a deep truth in this idea, as applied to a person's thinking.

The objects of a person's proper concern are multifarious, but they surely include her thinking. Thinking is how persons develop themselves into unique and robust members of a rational community. Consequently, it is an activity that each person must perform for themselves. To abandon one's thinking or to surrender it to another is to abandon or surrender one's very existence as a person.

Crucially though, the importance of thinking for oneself does not entail a need to think in isolation. To the contrary, developing oneself as a member of a rational community requires thinking with others, though the way in which an individual's thinking is joined to that of the community is of critical importance. The time, place, manner, and terms of such a joinder to communal thinking are sensitive matters, and so subject to various kinds of moral regulation.

Privacy is one such regulation. Its purpose is to ensure that thinking with others amplifies rather than stifles personal development, which it does by ordering solitary and communal thinking, so that persons have an opportunity to be concerned with their thinking, first by conducting it to the best of their ability, and then through exposure to the concern of others, in a community of thinking.

I argued that holding persons responsible essentially exhibits this kind of concern. In solitary thinking, persons show concern for themselves, by holding themselves responsible for what they

think – for thinking well – by attempting to have thoughts that are true, well-intentioned, creative, authentic, and so on. Others show concern in this way too. In holding someone responsible for her thinking, others show their concern for it (and their own), by embedding it within further, collective deliberation. The first step in such a discussion, and so the first step in holding someone responsible for their thinking is to know what they think. One shows care for someone by hearing them out. In doing this, one gets them right, and so lays the foundation for further engagement with them, as the person they are.

The problem that privacy addresses is that concern for the thinking of others can be excessive. By reading a person's diary or reading her mind, one shows concern for her thinking, but without her warranting that she is satisfied that the contents *are* her best thinking on the issue. The result is a misrepresentation of her at the level of what she thinks, and the interruption of her chance to set herself in order. Hence, accessing someone's thoughts by reading her mind or her diary is the most fundamental way in which one shows concern for her thinking in a way that wrongfully interferes with her opportunity to show that concern herself. "Mind your own business" captures this structure of the wrong. It captures the fact that violating a person's informational privacy interest in her mental activity consists in being concerned with, i.e., "minding," her business in a way that damages or hijacks her ability to conduct it herself.

This story about holding a person responsible for her thinking prematurely, and so wrongfully minding her business, covers just one kind of case, but it is positioned to extend to others. For instance, the most basic form of holding responsible consists in gathering information, but it may take other forms, such as various legal remedies or duties. If identifying a person's unexpressed thoughts wrongfully holds them responsible for them, then taxing them, or punishing them for those thoughts will be just as wrong, and arguably, more so, since they impose additional

burdens. Thought crimes will be ruled out, but so will thought taxes. Both are violations of mental privacy.

An outstanding question for this view of mental privacy is whether it generalizes to other kinds of privacy. I will close by sketching some support for thinking it does. The unifying idea is that of a person's identity, its public reality (it comes about through others ratifying it or interacting with it), and how forming it over time is importantly one's business.

The identity of a person has different facets. One aspect of it consists of her thoughts. I argued that thoughts are one of the most basic units of thinking, and so they are one's business in a deep way, though as I argued, not always to the exclusion of others.

A further aspect of a person's identity consists in what "labels" she accepts, where accepting a label is a kind of thought or mental maneuver that unifies and mobilizes other thoughts in a commitment-like way. This kind of acceptance is "identifying as." For instance, one may identify as Catholic or gay, where this is a matter of taking on certain commitments, thinking certain things, and having special relationships with certain others. What one identifies as has public significance and so is often how one wishes to be recognized, but preparing to adopt a label – and to ultimately, perhaps, come out as that sort of person – is unlike preparing to make a claim or state one's intentions. It is not done just by reflecting, but by other things too, like having experiences, and trying some of the activities and commitments that the label comes with. Since one's identity is one's business, one will have some privacy interests in the ways in which one prepares to adopt identities, such as by associating with others, and by having access to private spaces in which to do so.

A further aspect of one's identity will be facts about oneself that fall short of being an identity. Many records are like this. One's medical records, educational records, financial records, and criminal records, all characterize who one is. Privacy for these things depends on the role they

play in one's public identity. In a capitalist economic system, a person's financial records will be her business because they are the record of how she, literally, conducts her business. A person's health records will be her business in virtue of how health figures in the life of a person, which will be influenced by the health system in place. Health needs are a vulnerability and so are managed partly through privacy. The point is that records about a person will be her business to the degree that they reflect information about a sphere of activity that is her business.

A final aspect of one's identity is one's body. Not facts about it, as would be contained in a medical record, but the corporeality of it. Here, we are again engaged in experimentation that precedes publicization, though of a kind that is not as fundamental as that of making up our minds. Still, persons interact with their bodies in a struggle to accept them, become comfortable in them, and to use them in various activities with others. Thus, their premature revelation can be prejudicial. A common example is the naked body or persons engaged in sex, but such an interest may be implicated by all sorts of things, such as bed head, misapplied makeup, or broccoli in one's teeth. As a result, persons have a privacy interest in the time, place, and manner that they present their bodies to others, especially the public at large.

The picture that emerges, just as a sketch, is that of persons growing into themselves: into their personhood, into their social persona, and into their bodies, by making those very things the site of experimentation and testing, by changing what they think, changing what they identify as, and changing their bodies. These changes warrant privacy.

BIBLIOGRAPHY

12 Angry Men, Directed by Sidney Lumet, Orion Nova, 1957.

Abood v. Detroit Board of Education, 431 U.S. 209 (1977).

Adams, Robert. "Involuntary Sins." *Philosophical Review*, vol. 94, no. 2, 1985, pp. 3-31.

Allen, Ronald. "Theorizing About Self-Incrimination." *Cardozo Law Review*, vol. 30, no. 3, 2008, pp. 729-750.

Alter, Charlotte. "Former Auschwitz Guard Apologizes at Trial: 'I Am Ashamed.'" *Time,* 29 Apr. 2016, https://time.com/4312199/auschwitz-nazi-guard-apology-reinhold-hanning/.

Amar, Akhil Reed, and Renee B. Lettow. "Fifth Amendment First Principles: The Self-Incrimination Clause." *Michigan Law Review*, vol. 93, no. 5, 1995, pp. 857-928.

Amar, Akhil Reed. "Against Exclusion (Except to Protect Truth or Prevent Privacy Violations)." *Harvard Journal of Law and Public Policy*, vol. 20, no. 2, 1997, pp. 457-466.

American Psychiatric Association, *Principles of Medical Ethics with Annotations Especially Applicable to Psychiatry*, 2013.

Anonymous. *A Warning.* Twelve, 2019.

Arendt, Hannah. *The Human Condition.* Doubleday, 1959.

Arpaly, Nomy. "On Acting Rationally Against One's Best Judgment." *Ethics*, vol. 110, no. 3, 2000, pp. 488-513.

AT&T Communications, Inc. v. Wiltel, Inc., 1 F.3d 1201 (Fed. Cir. 1993).

Aviv, Rachel. "How Albert Woodfox Survived Solitary." *The New Yorker,* 8 Jan. 2017, https://www.newyorker.com/magazine/2017/01/16/how-albert-woodfox-survived-solitary

Balboa Island Village Inn v. Lemen, 156 P.3d 339 (Cal. 2007).

Baldwin, James. *The Fire Next Time.* 1963. Random House, 1993.

Bartholomew, Christine P. "Exorcising the Clergy Privilege." *Virginia Law Review*, vol. 103, no. 6, 2017, pp. 1015-1076.

Baumgarten, Elias. "Curiosity as Moral Virtue." *International Journal of Applied Philosophy*, vol. 15, no. 2, 2001, pp. 169-184.

Beauchamp, Scott. "Why Clinton's Iraq Apology Still Isn't Enough." *The Atlantic*, 8 Sept. 2016, https://www.theatlantic.com/international/archive/2016/09/clinton-iraq-bush-war-hussein-wmd-senate/499160/.

Beckerman, Gal. "What Do Famous People's Bookshelves Reveal?" *New York Times*, Updated 27 July 2020, https://www.nytimes.com/2020/04/30/books/celebrity-bookshelves-tv-coronavirus.html.

Beglin, David. "Two Strawsonian Strategies for Accounting for Morally Responsible Agency." *Philosophical Studies*, vol. 177, no. 8, 2020, pp. 2341-2364.

Berman, Mitchell N. "Attempts, in Language and in Law." *Jerusalem Review of Legal Studies*, vol 6, no 1, Dec. 2012, pp. 1-19.

Bester, Alfred. *The Demolished Man*. J. Boylston & Company, 2018.

Bilz, Kenworthey. "Self-Incrimination Doctrine is Dead; Long Live Self-Incrimination Doctrine: Confessions, Scientific Evidence, and the Anxieties of the Liberal State." *Cardozo Law Review*, vol. 30, no. 3, 2008, pp. 807-870.

Blackstone, William, et al. *Commentaries on the Laws of England.* Oxford University Press, 2016.

Blitz, Marc. "Constitutional Safeguards for Silent Experiments in Living: Libraries, the Right to Read, and a First Amendment Theory for an Unaccompanied Right to Receive Information." *University of Missouri-Kansas City Law Review*, vol. 74, no. 4, 2004, pp. 799-882.

---. "The Freedom of 3d Thought: The First Amendment in Virtual Reality." *Cardozo Law Review*, vol. 30, no. 3, 2008, pp. 1141-1244.

Boyle, Matthew. "Two Kinds of Self-Knowledge." *Philosophy and Phenomenological Research*, vol. 78, no. 1, 2009, pp. 133-164.

Brandenburg v. Ohio, 395 U.S. 444 (1969).

Bratman, Michael. *Intentions, Plans, and Practical Reason*. Harvard University Press, 1987.

Brennan-Marquez, Kiel. "A Modest Defense of Mind Reading." *Yale Journal of Law and Technology*, vol. 15, no. 1, 2013, pp. 214-272.

Brenner, Susan and Lori E. Shaw. *Federal Grand Jury: A Guide to Law & Practice*. 2nd ed., Thompson West, 2020.

Bromwich, David. "How Publicity Makes People Real." *Social Research*, vol. 68, no. 1, 2001, pp. 145-171.

Brooks, Peter. *Troubling Confessions*. University of Chicago Press, 2000.

Brudner, Alan. *Punishment and Freedom: A Liberal Theory of Penal Justice*. Oxford University Press, 2009.

Buckley v. American Constitutional Law Foundation, Inc., 525 U.S. 182 (1999).

Burge, Tyler. "Content Preservation." *The Philosophical Review*, vol. 102, no. 4, 1993, pp. 457-488.

Calhoun, Cheshire. "Standing for Something." *The Journal of Philosophy*, vol. 92, no. 5, 1995, pp. 235-260.

Calvert, Clay. "Freedom of Thought, Offensive Fantasies and the Fundamental Human Right to Hold Deviant Ideas: Why the Seventh Circuit Got It Wrong in Doe v. City of Lafayette, Indiana." *Pierce Law Review,* vol. 3, no. 2, 2005, pp. 125-160.

---. *Voyeur Nation: Media, Privacy, and Peering in Modern Culture*. Westview, 2000.

Cantwell v. State of Connecticut, 310 U.S. 296 (1940).

Catholic Church. *Catechism of the Catholic Church*. 2nd ed., Our Sunday Visitor, 2000.

Channel 10, Inc. v. Gunnarson, 337 F. Supp. 634 (D. Minn. 1972).

Chartier, Roger. "The Practical Impact of Writing." *A History of Private Life*. Translated by Arthur Goldhammer. Edited by Roger Chartier, vol. 3, Belknap Press, 1989, pp. 111-159.

Choi, Bryan H. "The Anonymous Internet." *Maryland Law Review*, vol. 72, no. 2, 2013, pp. 501-570.

Cohen, Julie E. "What Privacy is For." *Harvard Law Review*, vol. 126, no. 7, May 2013, pp. 1904-1933.

Cohen, Julie E. "A Right to Read Anonymously: A Closer Look at "Copyright Management" in Cyberspace." *Connecticut Law Review*, vol. 28, no. 4, 1996, pp. 981-1040.

Commonwealth v. Hardick, 380 A.2d 1235 (Pa. 1977).

Commonwealth v. Jones, No. 3345OF2001, 2002 WL 34401140 (Pa. Ct. Comm. Pl. Oct. 08, 2002).

Couch v. United States, 409 U.S. 322 (1973).

Cour d'appel [CA][regional court of appeal] Paris, Mar. 19, 1947, D.P. 20. (*L'affaire Rouault)*

Dan-Cohen, Meir. "Revising the Past: On the Metaphysics of Repentance, Forgiveness, and Pardon." *Forgiveness, Mercy, and Clemency*, edited by Austin Sarat and Nasser Hussain, Stanford University Press, 2007, pp. 117-137.

Davidson, Donald. "How is Weakness of the Will Possible?" *Essays on Actions and Events*, 2nd ed., Oxford, 2001, pp. 21-42.

De Melo-Martin, Inmaculada. "The Ethics of Anonymous Gamete Donation." *Hastings Center Report*, vol. Mar.-Apr. 2014, pp. 28-35.

Deutsch, Morton, and Harold B. Gerard. "A Study of Normative and Informational Social Influences Upon Individual Judgment." *Journal of Abnormal and Social Psychology*, vol. 51, no. 3, 1955, pp. 629-636.

Dick, Philip. *Minority Report and Other Classic Stories.* 1956. Citadel, 2016.

Dolinko, David. "Is there a Rationale for the Privilege Against Self-Incrimination?" *UCLA Law Review*, vol. 33, no. 4, 1986, pp. 1063-1148.

Dolinko, David. "Some Thoughts About Retributivism." *Ethics*, vol. 101, no. 3, 1991, pp. 537-559.

Dominus, Susan. "Snoopers on Subway, Beware Digital Books." *New York Times*, 31 Mar. 2008, https://www.nytimes.com/2008/03/31/nyregion/31bigcity.html.

Donohue, Jenna. *A Deliberative Conception of Complicity*. Forthcoming. University of California, Los Angeles, PhD dissertation.

Duff, R.A. "Risks, Culpability, and Criminal Liability." *Seeking Security: Pre-empting the Commission of Criminal Harms*, edited by G.R. Sullivan and Ian Dennis, Hart, 2012, pp. 121- 142.

---. *Answering for Crime*. 2007. Hart, 2009.

---. *Criminal Attempts*. Oxford University Press, 1996.

Dworkin, Gerald. "Punishment for Intentions." *Mind*, vol. 75, no. 299, 1966, pp. 396-404.

"Economist Media Directory." *The Economist*, https://mediadirectory.economist.com/. Accessed 15 Dec. 2020.

Epstein, Richard A. "Privacy, Property Rights, and Misrepresentations." *Georgia Law Review*, vol. 12, no. 3, Spring 1978, pp. 455-474.

Erez, Edna. "Thou Shalt not Execute, Hebrew Law Perspective on Capital Punishment." *Criminology*, vol. 19, no. 1, 1981, pp. 25-43.

Essert, Chris. "Property and Homelessness." *Philosophy and Public Affairs*, vol. 44, no. 4, 2016, pp. 266-295.

Fairfax, Robert. "Grand Jury Discretion and Constitutional Design." *Cornell Law Review*. vol. 93, no. 4, 2008, pp. 703-764.

Falls-Corbitt, Margaret, and Michael McLain. "God and Privacy." *Faith and Philosophy*, vol. 9, no. 3, 1992, pp. 369-386.

Farahany, Nina. "Incriminating Thoughts." *Stanford Law Review*, vol. 64, no. 2, 2012, pp. 351-408.

Faulders, Katherine. "'Anonymous,' Author of White House tell-all book, revealed to be Miles Taylor." *ABC News,* 28 Oct. 2020, https://abcnews.go.com/Politics/anonymous-author-white-house-book-revealed-miles-taylor/story?id=73884296.

Federal Rules of Evidence. The Legal Information Institute, 2020, https://www.law.cornell.edu/rules/fre. Accessed 15 Dec. 2020.

Forster, E.M. "Anonymity: An Inquiry." *Calendar of Modern Letters*, vol. 2, no. 9, 1925, pp. 145-156.

Fried, Charles. "Privacy [A Moral Analysis]." *Philosophical Dimensions of Privacy*, edited by Ferdinand Schoeman, Cambridge University Press, 1984, pp. 203-222.

---. "The New First Amendment Jurisprudence: A Threat to Liberty." *University of Chicago Law Review*, vol. 59, no. 1, Winter 1992, pp. 225-254.

Froomkin, Michael. "Lessons Learned Too Well: Anonymity in A Time of Surveillance." *Arizona Law Review,* vol. 59, no. 1, 2017, pp. 95-160.

Fulda, Joseph. "The Ethics of Pseudonymous Publication." *Journal of Information Ethics*, vol. 16, no. 2, 2007, pp. 75-89.

Gaetani v. Hadley, No. 14-30057-MGM, 2016 WL 593496 (D. Mass. Feb. 12, 2016).

Gardner, James. "Anonymity and Democratic Citizenship." *William and Mary Bill of Rights Journal*, vol. 19, no. 4, 2011, pp. 927-958.

Gavison, Ruth. "Privacy and the Limits of the Law." *Yale Law Journal*, vol. 89, no. 3, 1980, pp. 421-471.

Gerstein, Robert. "Privacy and Self-Incrimination." *Ethics*, vol. 80, no. 2, 1970, pp. 87-101.

---. "Punishment and Self-Incrimination." *American Journal of Jurisprudence*, vol. 16, no. 1, 1971, pp. 84-94.

---. "The Demise of Boyd: Self-Incrimination and Private Papers in the Burger Court." *UCLA Law Review*, vol. 27, no. 2, 1979, pp. 343-397.

Gertz v. Robert Welch, Inc., 418 U.S. 323 (1974).

Gover, K.E. *Art and Authority.* Oxford University Press, 2018.

Greely, Henry T. "Neuroscience, Mind-Reading, and the Courts: The Example of Pain." *Journal of Healthcare Law and Policy*, vol. 18, no. 2, 2015, pp. 171-206.

Green, Michael. "The Privilege's Last Stand: The Privilege Against Self-Incrimination and the Right to Rebel Against the State." *Brooklyn Law Review*, vol. 65, no. 3, 1999, pp. 627-716.

Gross, Larry. *Contested Closets: The Politics and Ethics of Outing.* University of Minnesota Press, 1993.

Halpern, Sue. *Migrations to Solitude.* First Vintage Books, 1992.

Hamilton, Alexander, et al. *The Federalist: With Letters of Brutus.* Edited by Terence Ball, Cambridge University Press, 2003.

Harrigan, Casey. "Against Dogmatism: A Continued Defense of Switch-Side Debate." *Contemporary Argumentation and Debate*, vol. 29, 2008, pp. 37-66.

Hausman, David. "How Congress Could Reduce Job Discrimination by Promoting Anonymous Hiring." *Stanford Law Revie*w, vol. 64, no. 5, 2012, pp. 1343-1370.

Hayward, Allison. "Bentham & Ballots: Tradeoffs Between Secrecy and Accountability in How We Vote." *Journal of Law and Politics*, vol. 26, no. 1, 2010, pp. 39-80.

Helmreich, Jeff. *More than Words: Stances as an Alternative Model for Apology, Forgiveness and Similar Speech Acts.* 2013. University of California, Los Angeles, PhD dissertation.

Herman, Barbara. "The Morality of Everyday Life." *Proceedings and Addresses of the American Philosophical Association*, vol. 74, no. 2, 2000, pp. 29-45.

Hesse, Herman. *Steppenwolf.* Translated by Joseph Mileck and Horst Frenz, Picador, 1963.

Hieronymi, Pamela. "Believing at Will." *Canadian Journal of Philosophy*, vol. 35, no. sup 1, 2009, pp. 149-187.

---. "Controlling Attitudes." *Pacific Philosophical Quarterly*, vol. 87, no. 2, 2006, pp. 45-74.

---. "Responsibility for Believing." *Synthese*, vol. 161, no. 3, 2008, pp. 357-373.

---. "The Use of Reasons in Thought (and the Use of Earmarks in Arguments)." *Ethics*, vol. 124, no. 1, 2013, pp. 114-127.

Hong, Nicole. "Ex-C.I.A. Analyst Faces Trial in Biggest Leak of Agency's History." *New York Times*, 4 Feb. 2020, https://www.nytimes.com/2020/02/04/nyregion/cia-leak-wikileaks-trial-Joshua-Schulte.html.

Hui, Sylvia. "J.K. Rowling Revealed as Writer of Crime Novel." *Washington Post*, 14 July 2013, https://wapo.st/2UEpxBm.

Hurley v. Irish-American Gay, Lesbian and Bisexual Group of Boston, Inc., 515 U.S. 557 (1995).

Husak, Douglas. "Does Criminal Liability Require an Act?" *Philosophy of Criminal Law: Selected Essays*, edited by Douglas Husak, Oxford University Press, 2010, pp. 17-52.

In re Establishment Inspection of Hern Iron Works, Inc., 881 F.2d 722 (9th Cir. 1989).

Inness, Julie. *Privacy, Intimacy, and Isolation.* Oxford University Press, 1992.

International Society for Krishna Consciousness of Berkeley, Inc. v. Kearnes, 454 F. Supp. 116 (E.D. Cal. 1978).

Jagodzinski, Cecile. *Privacy and Print.* University Press of Virginia, 1999.

John Doe v. Reed, 561 U.S. 186 (2010).

Kaminski, Margot. "Real Masks and Real Name Policies: Applying Anti-Mask Case Law to Anonymous Online Speech." *Fordham Intellectual Property, Media & Entertainment Law Journal*, vol. 23, no. 3, 2013, pp. 815-896.

Kaminski, Margot, and Shane Witnov. "The Conforming Effect: First Amendment Implications of Surveillance Beyond Chilling Speech." *University of Richmond Law Review*, vol. 49, no. 2, 2015, pp. 465-518.

Kastigar v. United States, 406 U.S. 441 (1972)

Kateb, George. "On Being Watched and Known." *Social Research*, vol. 68, no. 1, 2000, pp. 269-298.

Kendrick, Leslie. "Speech, Intent, and the Chilling Effect." *William & Mary Law Review,* vol. 54, no. 5, Apr. 2013, pp. 1633-1692.

Keyishian v. Bd. of Regents, 385 U.S. 589 (1967).

Kimpflen, John and Karl Oakes. *Corpus Juris Secundum Torts*. vol. 86, Thompson West, 2020.

King, Barry. "Stardom, Celebrity, and Para-Confession." *Social Semiotics*, vol. 18, no. 2, 2008, pp. 115-132.

Kleist, Chad. "Huck Finn the Inverse Akratic: Empathy and Justice." *Ethical Theory and Moral Practice*, vol. 12, no. 3, 2008, pp. 257-266.

Kramer, Susan. "'We Speak to God with Our Thoughts': Abelard and the Implications of Private Communication with God." *Church History*, vol. 69, no. 1, 2000, pp. 18-40.

Laden, Anthony. *Reasoning: A Social Picture.* Oxford University Press, 2012.

LaFave, Wayne R., *Substantive Criminal Law*. 3rd ed., Thompson West, 2017.

Landi, Ann. "When Is Artwork Finished?" *ARTNews*, February 24, 2014, https://www.artnews.com/art-news/news/when-is-an-artwork-finished-2383/.

Larson, Thomas. *The Memoir and the Memoirist.* Ohio University Press, 2007.

Ledewitz, Bruce. "Civil Disobedience, Injunctions, and the First Amendment." *Hofstra Law Review*, vol. 19, no. 1, 1990, pp. 67-142.

Lemley, Mark A., and Eugene Volokh. "Freedom of Speech and Injunctions in Intellectual Property Cases." *Duke Law Journal,* vol. 48, no. 2, Nov. 1998, pp. 147-242.

Lerner, Jennifer S., and Philip E. Tetlock. "Accounting for the Effects of Accountability." *Psychological Bulletin*, vol. 125, no. 2, 1999, pp. 255-275.

Lever, Annabelle. "Mill and the Secret Ballot: Beyond Coercion and Corruption." *Utilitas*, vol. 19, no. 3, 2007, pp. 354-378.

---. "Privacy and Democracy: What the Secret Ballot Reveals." *Law, Culture, and the Humanities*, vol. 11, no. 2, 2015, pp. 164-183.

Levmore, Saul. "The Anonymity Tool." *University of Pennsylvania Law Review*, vol. 144, no. 5, 1996, pp. 2191-2236.

Lidsky, Lyrissa, and Thomas F. Cotter. "Authorship, Audiences, and Anonymous Speech." *Notre Dame Law Review*, vol. 82, no. 4, 2007, pp. 1537-1604.

Lippke, Richard. "No Easy Way Out." *Law and Philosophy*, vol. 27, no. 4, 2008, pp. 383-414.

Livingston, Margit. "Disobedience and Contempt." *Washington Law Review,* vol. 75, no. 2, Apr. 2000, pp. 345-428.

Lodge, David. *Consciousness and the Novel.* Harvard University Press, 2002.

Louch, A.R. "Scientific Discovery and Legal Change." *The Monist*, vol. 49, no. 3, 1965, pp. 485-503.

Macklem, Timothy. *Independence of Mind.* Oxford University Press, 2008.

Malamud Smith, Janna. *Private Matters: In Defense of the Personal Life.* Seal Press, 1997.

Marmor, Andrei. "What Is the Right to Privacy?" *Philosophy & Public Affairs*, vol. 43, no. 1, 2015, pp. 3-26.

"McCain Counters Obama 'Arab' Question." *YouTube*, uploaded by the Associated Press, 11 Oct. 2008, https://www.youtube.com/watch?v=jrnRU3ocIH4.

McCloskey, H.J. "An Examination of Restricted Utilitarianism." *Philosophical Review*, vol. 66, no. 4, 1957, pp. 466-485.

McDaniel v. Pettigrew, 536 S.W.2d 611 (Tex. Civ. App. 1976).

McIntyre v. Ohio Elections Commission 514 U.S. 334 (1995).

McKay, Robert. "Self-Incrimination and the New Privacy." *Supreme Court Review*, vol. 1967, 1967, pp. 193-232.

Meilaender, Gilbert. "Why Remember?" *First Things*, August 2003, https://www.firstthings.com/article/2003/08/why-remember. Accessed 15 Dec. 2020.

Menand, Louis. "Lives of Others." *The New Yorker* 25 Dec. 1969, https://www.newyorker.com/magazine/2007-08ds/lives-of-others.

Mendlow, Gabriel S. "Why Is It Wrong to Punish Thought?" *Yale Law Journal,* vol. 127, no. 8, June 2018, pp. 2342-2387.

Miami Herald Publishing Co. v. Tornillo, 418 U.S. 241 (1974).

Michelson v. United States, 335 U.S. 469 (1948).

Mill, John Stuart. *Considerations on Representative Government,* Henry Holt and Company, 1873. Google Books, https://www.google.com/books/edition/Considerations_on_Representative_Governm/16FFAQAAMAAJ?hl=en&gbpv=1.

---. *On Liberty*, Penguin, 1985.

"Model Rules of Professional Conduct." *The American Bar Association*, 2020, https://www.americanbar.org/groups/professional_responsibility/publications/model_rules_of_professional_conduct/model_rules_of_professional_conduct_table_of_contents/. Accessed 15 Dec. 2020.

Moran, Richard. "Getting Told and Being Believed." *Philosophers' Imprint*, vol. 5, 2005, pp. 1-29.

---. "Problems of Sincerity." *Proceedings of the Aristotelian Society*, vol. 105, no. 3, 2005, pp. 341-361.

---. *The Exchange of Words.* Oxford University Press, 2018.

Murphy v. Waterfront Commission of NY Harbor, 378 U.S. 52 (1964).

NAACP v. Alabama Ex Rel. Patterson, 357 U.S. 449 (1958).

National Commission on Terrorist Attacks. *The 9/11 Commission Report.* WW Norton, 2004.

National Federation of Independent Business v. Sebelius, 567 U.S. 519 (2012).

National Institute of Family & Life Advocates v. Becerra, 138 S. Ct. 2361 (2018).

Navarette v. California, 572 U.S. 393 (2014).

Near v. Minnesota, 283 U.S. 697 (1931).

Nehamas, Alexander. "An Essay on Beauty and Judgment." *The Threepenny Review,* vol. 80, Winter 2000.

Newman, Jon. "Copyright Law and the Protection of Privacy." *Columbia-VLA Journal of Law & Arts*, vol. 12, no. 4, 1988, pp. 459-480.

Nissenbaum, Helen. *Privacy in Context.* Stanford Law Books, 2010.

---. "Privacy as Contextual Integrity." *Washington Law Review,* vol. 79, no. 1, Feb. 2004, pp. 119-158.

Nygaard, Richard Lowell. "The Maligned Per Curiam: A Fresh Look at an Old Colleague." *Scribes Journal of Legal Writing.* vol. 5, 1995, pp. 41-50.

Oatley, Keith, and Maja Djikic. "Writing as Thinking." *Review of General Psychology*, vol. 12, no. 1, 2008, pp. 9-27.

Oatley, Keith. "Thinking Deeply in Reading and Writing." *The Edge of the Precipice*, edited by Paul Socken, McGill-Queen's University Press, 2013, pp. 175-191.

Ohralik v. Ohio State Bar Association, 436 U.S. 447 (1978).

Pardo, Michael. "Disentangling the Fourth Amendment and the Self-Incrimination Clause." *Iowa Law Review*, vol. 90, no. 5, 2005, pp. 1857-1904.

---. "Self-Incrimination and the Epistemology of Testimony." *Cardozo Law Review*, vol. 30, no. 3, 2008, pp. 1023-1046.

Peterson, Andrea. "How Washington's Last Remaining Video Rental Store Changed the Course of Privacy Law." *Washington Post*, 28 Apr. 2014, https://www.washingtonpost.com/news/the-switch/wp/2014/04/28/how-washingtons-last-remaining-video-rental-store-changed-the-course-of-privacy-law/.

Picinali, Federico. "A Retributive Justification for Not Punishing Bare Intentions or: On the Moral Relevance of the 'Now-Belief'." *Law and Philosophy*, vol. 32, no. 4, 2013, pp. 385-403.

Pitkin, Hanna. *The Concept of Representation*. University of California Press, 1967.

Ponesse, Julie. "The Ties that Blind: Conceptualizing Anonymity." *Journal of Social Philosophy*, vol. 45, no. 3, 2014, pp. 304-322.

Posner, Eric, and Adrian Vermeule. "Legislative Entrenchment: A Reappraisal." *Yale Law Journal*, vol. 111, no. 7, 2002, pp. 1665-1705.

Post, Robert C. "Three Concepts of Privacy." *Georgetown Law Journal*, vol. 89, no. 6, June 2001, pp. 2087-2098.

Poulet, Georges. "Phenomenology of Reading." *New Literary History*, vol. 1, no. 1, 1969, pp. 53-68.

Quantity of Copies of Books v. State of Kansas, 378 U.S. 205 (1964).

Rachels, James. "Why Privacy is Important." *Philosophical Dimensions of Privacy*, edited by Ferdinand Schoeman, Cambridge University Press, 1984, pp. 290-299.

Radzik, Linda. "Tort Processes and Relational Repair." *Philosophical Foundations of the Law of Torts*, edited by John Oberdiek, Oxford University Press, 2014, pp. 231- 249.

Rakoff, Todd. "The Implied Terms of Contracts: Of 'Default Rules' and 'Situation-Sense'." *Good Faith and Fault in Contract Law*, edited by Jack Beatson and Daniel Friedmann, Oxford University Press, 1997, pp. 191-228.

Re, Richard M. "Beyond the Marks Rule." *Harvard Law Review*, vol. 132, no. 7, 2019, pp. 1943-2008.

Redish, Martin. "The Proper Role of the Prior Restraint Doctrine in First Amendment Theory." *Virginia Law Review*, vol. 70, no. 1, 1984, pp. 53-100.

Register, Lisa. "The Phenomenology of Intimacy." *Journal of Social and Personal Relationships*, vol. 9, 1992, pp. 467-481.

Rich, Michael L. "Limits on the Perfect Preventive State." *Connecticut Law Review,* vol. 46, no. 3, Feb. 2014, pp. 883-936.

Richards, David A. J. "Constitutional Legitimacy, the Principle of Free Speech, and the Politics of Identity." *Chicago-Kent Law Review,* vol. 74, no. 2, 1999, pp. 779-824.

Richards, Neil. "The Perils of Social Reading." *Georgetown Law Review*, vol. 101, no. 3, 2013, pp. 689-724.

---. *Intellectual Privacy.* Oxford University Press, 2015.

Robbins, Ira P. "Hiding Behind the Cloak of Invisibility: The Supreme Court and Per Curiam Opinions." *Tulane Law Review*, vol. 86, no. 6, 2012, pp. 1197-1242.

Roberts v. United States, 445 U.S. 552 (1980).

Robinson v. California, 370 U.S. 660 (1962).

Roe v. Department of Defense, 947 F.3d 207 (4th Cir. 2020).

Rogowsky, Scott, "Taking Fake Book Covers on the Subway." *YouTube*, uploaded by the Chortle, 6 Apr. 2016, https://www.youtube.com/watch?v=jFxu9dOO4zk&feature=youtu.be.

Rosen, Jeffrey. *The Unwanted Gaze.* Random House, 2000.

Rubenfeld, Jed. "The Freedom of Imagination." *Yale Law Journal*, vol. 112, no. 1, 2002, pp. 1-60.

---. "The Right to Privacy." *Harvard Law Review*, vol. 102, no. 4, 1989, pp. 737-807.

S.E.C. v. Carriba Air, Inc., 681 F.2d 1318 (11th Cir. 1982).

Sarraute, Raymond. "Current Theory on the Moral Right of Authors and Artists Under French Law." *American Journal of Comparative Law*, vol. 16, no. 4, 1968, pp. 465-486.

Schedler, Andreas. "Conceptualizing Accountability." *The Self Restraining State: Power and Accountability in New Democracies*, edited by Andreas Schedler, Larry Diamond, and Marc Plattner, Lynne Rienner, 1999, pp. 13-28.

Schulhofer, Stephen. "Some Kind Words for the Privilege Against Self-Incrimination." *Valparaiso University Law Review*, vol. 26, no. 1, 1991, pp. 311-336.

Schwartz, Alexandra. "The 'Unmasking' of Elena Ferrante." *The New Yorker*, 3 Oct. 2016, https://www.newyorker.com/culture/cultural-comment/the-unmasking-of-elena-ferrante

Schwartz, Joanna C. "The Case against Qualified Immunity." *Notre Dame Law Review,* vol. 93, no. 5, May 2018, pp. 1797-1852.

Sealed Plaintiff v. Sealed Defendant, 537 F.3d 185 (2d Cir. 2008).

"See What We've Learned." *New York Times Magazine*, https://www.nytimes.com/interactive/2020/05/19/magazine/covid-quarantine-dust.html.

Settlemeyer, Carl, III. "Between Thought and Possession: Artists' 'Moral Rights' and Public Access to Creative Works." *Georgetown Law Journal*, vol. 81, no. 6, 1993, pp. 2291-2344.

Shapiro, Eugene L. "Examining an Underdeveloped Constitutional Standard: Trial in Absentia and the Relinquishment of a Criminal Defendant's Right to Be Present." *Marquette Law Review*, vol. 96, no. 2, 2012, pp. 591-622.

Sharp, Ann, and Laurance Splitter. *Teaching for Better Thinking.* Australian Council for Educational Research, 1995.

Shiffrin, Seana Valentine, "A Thinker Based Approach to Freedom of Speech." *Constitutional Commentary*, vol. 27, no. 2, 2011, 283-307.

---. "Caution About Character Ideals and Capital Punishment: A Reply to Sorell." *Criminal Justice Ethics*, vol. 21, no. 2, 2002, pp. 35-39.

---. "Paternalism, Unconscionability Doctrine, and Accommodation." *Philosophy and Public Affairs*, vol. 29, no. 3, 2000, pp. 205-250.

---. "What is Really Wrong with Compelled Association?" *Northwestern University Law Review*, vol. 99, no. 2, 2005, pp. 839-888.

---. *Speech Matters.* Princeton University Press, 2014.

Sidis v. FR Publishing Corporation, 113 F.2d 806 (1940).

Silverberg, Robert. *Dying Inside.* Tom Doherty, 2009.

Skopek, Jeffrey. "Anonymity, the Production of Goods, and Institutional Design." *Fordham Law Review*, vol. 82, no. 4, 2014, pp. 1751-1810.

---. "Reasonable Expectations of Anonymity." *Virginia Law Review*, vol. 101, no. 3, 2015, pp. 691-762.

Smith Ekstrand, Victoria, "The Many Masks of Anon: Anonymity as Cultural Practice and Reflections in Case Law." *Journal of Technology Law and Policy*, vol. 18, no. 1, 2013, pp. 1-36.

Smith v. Mann, 173 F.3d 73 (2d Cir. 1999).

Smith, Kerri. "Brain Decoding: Reading Minds." *Nature*, vol. 502, no. 7472, 2013, pp. 428-430.

Snead, O. Carter. "Memory and Punishment." *Vanderbilt Law Review*, vol. 64, no. 4, 2011, pp. 1195-1264.

Solove, Daniel. *Understanding Privacy.* Harvard University Press, 2008.

Souto-Manning, Mariana. "Challenging the Text and Context of (Re)Naming Immigrant Children: Children's Literature as Tools for Change." *Promoting Social Justice for Young Children*, edited by Beatrice S. Fennimore & A. Lin Goodwin, Springer, 2011, pp. 111-124.

Speiser v. Randall, 357 U.S. 513 (1958).

Stalnaker, Robert. "Common Ground." *Linguistics and Philosophy*, vol. 25, no. 5-6, 2002, pp. 701-721.

Stanley v. Georgia, 394 U.S. 557 (1969).

Stern, Stephanie. "The Inviolate Home: Housing Exceptionalism in the Fourth Amendment." *Cornell Law Review*, vol. 95, no. 5, 2010, pp. 905- 956.

Storr, Anthony. *Solitude: A Return to the Self.* Simon and Schuster, 1988.

Strauss, David. "Persuasion, Autonomy, and Freedom of Expression." *Columbia Law Review*, vol. 91, no. 2, 1991, pp. 334-371.

Stuntz, William. "Self-Incrimination and Excuse." *Columbia Law Review*, vol. 88, no. 6, 1988, pp. 1227-1296.

Sturgis, Daniel. "Is Voting a Private Matter?" *Journal of Social Philosophy*, vol. 36, no. 1, 2005, pp. 18-30.

Talley v. California, 362 U.S. 60 (1960).

Tenenbaum, Sergio. "Akrasia and Irrationality." *A Companion to the Philosophy of Action*, edited by Sandis O'Connor, Blackwell, 2010, pp. 274-282.

Theuns, Tom. "Jeremy Bentham, John Stuart Mill and the Secret Ballot: Insights from Nineteenth Century Democratic Theory." *Australian Journal of Politics and History*, vol. 63, no. 4, 2017, pp. 493-507.

Traynor, Roger. "Some Open Questions on the Work of State Appellate Courts." *University of Chicago Law Review*, vol. 24, no. 2, 1957, pp. 211-224.

Tribe, Laurence. "The Mystery of Motive, Private and Public: Some Notes Inspired by the Problems of Hate Crimes and Animal Sacrifice." *Supreme Court Review*, vol. 1993, 1993, pp. 1-36.

Tsai, George. "Rational Persuasion as Paternalism." *Philosophy and Public Affairs*, vol. 42, no. 1, 2014, pp. 78-112.

United States v. $11,500.00 in United States Currency, 869 F.3d 1062 (9th Cir. 2017).

United States v. El-Mezain, 664 F.3d 467 (5th Cir. 2011).

United States v. Guest, 383 U.S. 745 (1966).

United States v. Hubbell, 530 U.S. 27 (2000).

United States v. Karo, 468 U.S. 705 (1984).

United States v. Mansoori, 304 F.3d 635 (7th Cir. 2002).

United States v. Moccia, 681 F.2d 61 (1st Cir. 1982).

United States v. Resendiz-Ponce, 549 U.S. 102 (2007).

United States v. Santee Sioux Tribe of Nebraska, 135 F.3d 558 (8th Cir. 1998).

United States v. Shabani, 513 U.S. 10 (1994).

United States v. Soto-Barraza, No. 15-10586, 2020 WL 253560, at *2 (9th Cir. Jan. 17, 2020).

United States v. W. T. Grant Co., 345 U.S. 629 (1953).

United States v. White, 610 F.3d 956 (7th Cir. 2010).

United States v. Williams, 553 U.S. 285 (2008).

Uviller, H. Richard. "Evidence from the Mind of the Criminal Suspect: A Reconsideration of the Current Rules of Access and Restraint." *Columbia Law Review*, vol. 87, no. 6, 1987, pp. 1137-1212.

Vandamme, Pierre-Etienne. "Voting Secrecy and the Right to Justification." *Constellations*, vol. 25, no. 3, 2017, pp. 388-405.

Volokh, Eugene. "Crime-Facilitating Speech." *Stanford Law Review*. vol. 57, no. 4, 2005, pp. 1095-1222.

Vrasic v. Leibel, 106 So. 3d 485 (Fla. Dist. Ct. App. 2013).

Walker, Kent. "Where Everybody Knows Your Name: A Pragmatic Look at the Costs of Privacy and the Benefits of Information Exchange." *Stanford Technology Law Review*, 2000, 2000, pp. 1-50.

Warner, Michael. "Publics and Counterpublics." *Public Culture*, vol. 14, no. 1, 2002, pp. 49-90.

Wasby, Stephen L. "The Per Curiam Opinion: Its Nature and Functions." *Judicature*. Vol. 76, no. 1, 1992, pp. 29-38.

Wasserstrom, Richard. "Privacy: Some Arguments and Assumptions." *Philosophical Dimensions of Privacy*, edited by Ferdinand Schoeman, Cambridge University Press, 1984, pp. 317-323.

Watchtower Bible & Tract Society of New York, Inc. v. Village of Stratton, 536 U.S. 150 (2002).

Waterman, Megan Elyse. "Indorsing Infant Immunity: An Argument for Criminalizing Parents' Refusal to Immunize Their Children." Tulsa Law Review, vol. 51, no. 1, Summer 2015, pp. 153-180.

We Live in Public. Directed by Ondi Timoner, Interloper Films, 2009.

West Virginia Code Annotated § 3-1-4 (West).

West Virginia State Board of Education v. Barnette, 319 U.S. 624 (1943).

Whitman, James Q. "The Two Western Cultures of Privacy." *Yale Law Journal*, vol. 113, no. 6, 2004, pp. 1151-1222.

Whitney v. California, 274 U.S. 357 (1927).

"Why are the Economist's Writers Anonymous?", *The Economist,* 27 Mar. 2017, https://medium.economist.com/why-are-the-economists-writers-anonymous-8f573745631d. Accessed 15 Dec. 2020.

Yang, Jack. "Professor X's Notice and Choice Privacy Dilemma." *The Legal Geeks*, 16 Aug. 2016, http://thelegalgeeks.com/2016/08/16/professor-xs-notice-and-choice-privacy-dilemma/. Accessed 15 Dec. 2020.

Young, Meredith. "The Utility of Failure: A Taxonomy for Research and Scholarship." *Perspectives on Medical Education*, vol. 8, 2019, pp. 365-371.

Young v. State, 303 Md. 298 (1985).