Untangling Privacy: Losses Versus Violations

Jeffrey M. Skopek[*]

ABSTRACT: Increasingly powerful data mining and analysis technologies are being used to learn information and make decisions about people across all areas of life—ranging from employment and policing, to housing and health insurance—and it is widely thought that the key problems with this are privacy-related. There is also an emerging consensus in the literature that privacy rights lack a unified core. This Article demonstrates that these are both mistaken conclusions that derive from the conflation of privacy losses and violations, and it develops a theory of privacy that untangles these misunderstood concepts at the heart of privacy law. In clarifying the outcome-based criteria for privacy losses and their relationship with the path-based criteria for privacy violations, this theory provides value across two domains. First, regarding the coherence of the law, it demonstrates that a unified theory of privacy rights is possible despite significant disagreement about their content. Second, regarding the law's content, it challenges orthodox views about how the aggregation, use, and inference of personal information violate privacy rights.

[*] University of Cambridge Faculty of Law. J.D., Harvard Law School; Ph.D. (History and Philosophy of Science), University of Cambridge; A.B., Stanford University. For helpful conversations and comments at various stages of this project, I wish to thank Jennifer Anderson, Lionel Bently, Glenn Cohen, David Erdos, Urs Gasser, Stephen John, Greg Keating, Kathy Liddell, Tim Lewens, John Murphy, Nicholson Price, and everyone who provided comments when I presented this in workshops at the Department of History and Philosophy of Science and the Centre for Intellectual Property and Information Law at the University of Cambridge; the Faculty of Law at Academia Sinica, Taiwan; and at the following conferences: The Methodology and Ethics of Targeting (University of Cambridge); Privacy, Data Protection and Data-Sharing (University of Hong Kong); Legal Dimensions of Big Data in the Health and Life Sciences (University of Copenhagen); Policy-Making in the Big Data Era (University of Cambridge); Biodata World Congress 2016 (Wellcome Genome Campus); and Big Data, Health Law, and Bioethics (Harvard Law School).

I. INTRODUCTION

It is widely thought that the core problems posed by new technologies of personal data mining and analysis, as well as their solutions, can be explained in terms of privacy. Take, for example, the uses of personal data in these cases:

- an employer rejects a job applicant on the basis of a health trait inferred from non-health data in his application;
- a landlord screens out applicants on the basis of a proxy for religion;
- the police aggregate data about a person's public movements, thereby discovering the person's sexual and political orientations;
- an internet platform infers private facts about a person from his browsing history and uses this to tailor content;
- a government agency makes a decision about an individual entitlement on the basis of an algorithmic assessment that it cannot explain.

These and other related uses of personal data are widely seen as violating privacy rights.[1] This is often a mistaken diagnosis, however, which arises from a failure to differentiate between privacy losses and privacy violations. To understand and address the actual threats posed by new ways of accessing and using personal data, it is necessary to step back and clarify what privacy is —and what it is not.

For as long as privacy has been the subject of academic study, privacy scholars have highlighted that it is an ill-defined concept,[2] and for as long as they have tried to clarify it, their definitions have been rejected by others as being too broad, too narrow, or both.[3] In light of this history, Dan Solove has championed the growing view that we should abandon our attempt "to locate

1. *See, e.g.*, Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 95–106 (2014); Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477, 505–16, 518–21 (2006) [hereinafter Solove, *A Taxonomy of Privacy*].

2. *See, e.g.*, DANIEL J. SOLOVE, UNDERSTANDING PRIVACY 1 (2008) ("Privacy . . . is a concept in disarray. Nobody can articulate what it means."); Tom Gerety, *Redefining Privacy*, 12 HARV. C.R.-C.L. L. REV. 233, 233 (1977) ("Privacy is a legal wall badly in need of mending."); Jeffery L. Johnson, *Privacy and the Judgment of Others*, 23 J. VALUE INQUIRY 157, 157 (1989) (comparing the concept of privacy to "a haystack in a hurricane").

3. In 1978, David O'Brien concluded that the unitary definitions of privacy that had been developed by others were either "imprecise, or too broad, or too narrow." David M. O'Brien, *Privacy and the Right of Access: Purposes and Paradoxes of Information Control*, 30 ADMIN. L. REV. 45, 62 (1978). Nearly 25 years later, Dan Solove reached the same conclusion: "The most prevalent problem with the conceptions is that they are either too narrow or too broad. . . . Often, the same conceptions can suffer from being both too narrow and too broad." Daniel J. Solove, *Conceptualizing Privacy*, 90 CALIF. L. REV. 1087, 1094 (2002) [hereinafter Solove, *Conceptualizing Privacy*].

the 'essential' or 'core' characteristics of privacy," which he identifies as the cause of the deep disagreements in the literature.[4] The way forward, he argues, is to understand privacy as a Wittgensteinian "family resemblance" concept that is not unified by any essential features, but rather held together by a common pool of similar features.

While Solove's analysis is illuminating, it has hidden costs, which are avoided by an alternative diagnosis of the problems in the literature. As this Article will demonstrate, the two different trends that I have identified thus far—the reliance on privacy to articulate an ever-growing list of concerns about data-driven technologies, and the growing skepticism about whether privacy has a unifying core—both derive from the same error. They derive from the failure to differentiate between descriptive and normative accounts of privacy and the different questions that they answer (explored in Part II).

As used in ordinary language, the term privacy may refer either to a state of affairs (e.g., "the choice of glass walls for his new office reduced his privacy") or to a right (e.g., "the wiretap of his phone violated his privacy"). Claims about the state of affairs are value-neutral, whereas claims about the right are not. In having these two dimensions, privacy is similar to many other important moral and legal concepts, such as liberty and discrimination, where the two dimensions are generally recognized.[5] With privacy, however, the distinction, which was once widely recognized as significant, is now generally overlooked.[6] As a result, two questions have become conflated: the first is the descriptive question of whether a person has suffered a privacy loss (i.e., a loss of privacy as a state of affairs); the second is the normative question of whether

4. SOLOVE, *supra* note 2, at 8; *see also* JUDITH WAGNER DECEW, IN PURSUIT OF PRIVACY: LAW, ETHICS, AND THE RISE OF TECHNOLOGY 61 (1997) (arguing that "it is not possible to give a unique, unitary definition of privacy that covers all the diverse privacy interests" and that we should understand "privacy as a broad and multifaceted cluster concept"); Scott A. Anderson, *Privacy Without the Right to Privacy*, 91 MONIST 81, 82 (2008) (identifying reasons to be skeptical about a unified account of privacy and arguing instead for a "piecemeal approach to privacy"); David E. Pozen, *Privacy-Privacy Tradeoffs*, 83 U. CHI. L. REV. 221, 225–28 (2016) (discussing the "pluralistic turn" in privacy scholarship).

5. For example, as Dworkin explains: "We use 'liberty' in its flat sense simply to indicate the absence of constraint. . . . We use 'liberty' in its normative sense, on the other hand, to describe the ways in which we believe people ought to be free." RONALD DWORKIN, SOVEREIGN VIRTUE: THE THEORY AND PRACTICE OF EQUALITY 125 (2002); *see also* Ralf M. Bader, *Moralized Conceptions of Liberty*, *in* THE OXFORD HANDBOOK OF FREEDOM 59, 59–60 (David Schmidtz & Carmen E. Pavel eds., 2018) (identifying and discussing the difference between descriptive and normative conceptions of liberty).

6. For example, the distinction was once highlighted by Ferdinand Schoeman. *See* Ferdinand Schoeman, *Privacy: Philosophical Dimensions of the Literature*, *in* PHILOSOPHICAL DIMENSIONS OF PRIVACY: AN ANTHOLOGY 1, 3 (Ferdinand D. Schoeman ed., 1984); *see also* Hyman Gross, *The Concept of Privacy*, 42 N.Y.U. L. REV. 34, 35–36 (1967); O'Brien, *supra* note 3, at 75; W.A. Parent, *Privacy, Morality, and the Law*, 12 PHIL. & PUB. AFF. 269, 269–75 (1983). It is now often overlooked, or if acknowledged, dismissed as unimportant. *See, e.g.*, Adam Moore, *Defining Privacy*, 39 J. SOC. PHIL. 411, 425 (2008). But there are exceptions, for example, see Madison Powers, *A Cognitive Access Definition of Privacy*, 15 LAW & PHIL. 369, 370–72 (1996).

a person has suffered a privacy violation (i.e., a violation of privacy as a legal or moral right).

The conflation of privacy losses and violations is problematic for various reasons, including that it has generated significant misconceptions in the privacy literature (demonstrated in Part III). First, it has generated mistargeted critiques in which descriptive accounts of privacy are rejected for failing to answer normative questions. This type of analytical error, known in philosophy as a "category error," has led many to believe that there is greater disagreement in the literature than is actually the case. Second, the conflation of privacy losses and violations has generated misguided skepticism about the possibility of developing a unified theory of privacy rights. This skepticism assumes that a unified theory of privacy must be grounded in normative coherence (i.e., in an agreement about what constitutes a privacy violation). An alternative approach, which has been overlooked, is to ground it in descriptive coherence (i.e., in an agreement about what constitutes a privacy loss).

On the question of what constitutes a privacy loss, there is also disagreement in the literature, but shared judgements are easier to generate. Critical reflection on how the concept is used in ordinary language reveals that it is best understood as being defined by three criteria. On this account (developed in Part IV), a privacy loss occurs when one's personal information is accessed by another,[7] the means of access have epistemic merit,[8] and the information is true.[9]

It is worth highlighting that I use the term "information" broadly in these criteria, capturing cases of access that some see as non-informational.[10] (Cases

7. As Section IV.A argues, mere accessibility, lack of control over access, and access by non-persons might seem to constitute privacy losses, but closer analysis reveals that these criteria capture related but distinct matters.

8. As Section IV.B argues, the criterion of "knowledge" that has been assumed in many definitions of privacy is not in fact necessary, but at the other end of the epistemic spectrum, a mere lucky guess is insufficient.

9. As Section IV.C argues, access to false information can cause just as much harm as access to true information, but this is not a sufficient basis to classify it as a privacy loss.

10. For example, Anita Allen argues that privacy losses can arise from three distinct forms of accessibility: physical, dispositional, and informational. ANITA L. ALLEN, UNEASY ACCESS: PRIVACY FOR WOMEN IN A FREE SOCIETY 16–17 (1988); *see also* Ruth Gavison, *Privacy and the Limits of Law*, 89 YALE L.J. 421, 432–33 (1980) (identifying attention paid to an individual and physical access to an individual as non-informational privacy losses); Richard B. Parker, *A Definition of Privacy*, 27 RUTGERS L. REV. 275, 281 (1974) (arguing that what is gained from the five senses "could be called 'information,' but it would be misleading to use that word"). It seems to me that these "non-informational" cases of privacy loss can in fact be properly explained in terms of informational access, but my argument does not require that I take a position on this question. It is also worth highlighting here that my account avoids the common objection to information-based accounts of privacy that define loss in terms of access to *new* information. *See, e.g.*, DECEW, *supra* note 4, at 34; JULIE C. INNESS, PRIVACY, INTIMACY, AND ISOLATION 63–64 (1992); Gavison, *supra*, at 431–32.

of pure "decisional interference," on the other hand, are excluded.)[11] Note also that privacy is not a binary or categorical state of affairs, so the term "privacy loss" refers to a loss with respect to a given piece of information and with respect to the particular person or persons who accessed it, not to a total loss of privacy.

Once the criteria that define privacy losses are clarified, it becomes easier to identify the defining feature of privacy rights—they restrict the means of access by which privacy losses can permissibly occur. On this account (developed in Part V), a person suffers a privacy violation when a restriction on the permissible means of generating this type of access is breached. Thus, a key difference between privacy losses and violations is that losses are outcome-based, whereas violations are path-based. Privacy rights do not protect a reasonable expectation that privacy will be maintained, but rather a reasonable expectation that privacy will not be lost in certain ways.

For the avoidance of confusion, it is worth highlighting that although this theory of privacy defines violations in terms of losses (which provides for the coherence of privacy), it does not suggest that a privacy loss is a necessary element of a privacy violation. On the contrary, in line with common intuitions, it explains how a privacy violation can occur without the occurrence of a privacy loss.[12] For example, the police can violate your privacy rights by installing an unauthorized wiretap on your phone (breaching a restriction on a means of access), even if you do not end up speaking on the phone (so access is not obtained).

It is also worth highlighting that this theory does not take a position on the question of which means of access violate privacy rights, but rather provides a foundation for a wide range of positions on this and other related questions (such as why privacy is valuable and what types of information and spaces should be protected). Because it unifies privacy along its descriptive rather than its normative dimension, it is compatible with disagreement on these questions. For example, it is compatible with Helen Nissenbaum's argument that privacy rights should be understood as rights to "contextual

11. While some privacy scholars suggest that "decisional interference"—i.e., government interference with personal decisions regarding certain areas of life (e.g., contraception, abortion, intimate activity, etc.)—implicates privacy regardless of informational access, it is widely agreed that these restrictions are better described as losses of autonomy. *See, e.g.*, LAURENCE H. TRIBE, AMERICAN CONSTITUTIONAL LAW 1352 (2d ed. 1988); Louis Henkin, *Privacy and Autonomy*, 74 COLUM. L. REV. 1410, 1410–11 (1974); Gavison, *supra* note 10, at 438–39; Gross, *supra* note 6, at 38. However, many of these cases are not *pure* decisional interference cases, as they also involve informational access, as is discussed by Neil M. Richards, *The Information Privacy Law Project*, 94 GEO. L.J. 1087, 1106–16 (2006).

12. *Cf.* Schoeman, *supra* note 6, at 4 ("We can also envision situations in which we would want to say that a person has not in fact suffered loss of privacy but has suffered a violation of his right to privacy.").

integrity,"[13] as well as Lior Strahilevitz's argument that privacy rights should focus on protecting interests against intrusion and disclosure.[14]

At the same time, however, this theory of privacy has a critical edge,[15] challenging widespread claims about whether and how privacy rights are violated by data aggregation, the unconsented use of personal data, and the inference of private facts from disclosed data. Paying attention to the loss/violation distinction reveals that the scholarship on these issues has misinterpreted key Supreme Court cases, including the landmark technology cases of *Carpenter v. United States*[16] and *Kyllo v. United States*.[17] In addition, it helps clarify the normative reasons why privacy rights should not be expanded in the some of the ways that have been suggested.

To be clear, in challenging widespread claims about how the aggregation, unconsented use, and inference of personal information violate privacy rights, I am not arguing that these practices do not violate *any* rights or that restrictions on them are unjustified. Rather, my argument is that if restrictions are justified, the justifications must often be found outside what are properly regarded as privacy rights. It is a mistake to think that all wrongful uses of personal information implicate privacy.[18]

Further, even when privacy is at issue, it important to understand the relationship between privacy losses, privacy violations, and the harms that they entail or generate (often referred to loosely as "privacy harms"). Because both privacy losses and violations cause privacy harms, it is a mistake to think that the existence of a such a harm means that a privacy right has been violated.

Increased analytical precision on these matters will likely reveal that some complaints that have been characterized as privacy violations should not be legally actionable at all. In other situations, it may help us develop new protections that are justified by the actual interests at stake, such as autonomy and fairness (which are often conflated with privacy). However, increased precision here may also uncover previously unrecognized limits; for when

13. HELEN NISSENBAUM, PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE 9, 127 (2010). On Nissenbaum's theory of privacy as contextual integrity, the right to privacy is "[the] right to *appropriate* flow of personal information[,]" and inappropriate information flows are those that violate context specific informational norms, which differ depending on the type of information at issue, the actors involved, and the principles under which the information is transmitted. *Id.* at 9, 127, 143.

14. *See generally* Lior Jacob Strahilevitz, *Reunifying Privacy Law*, 98 CALIF. L. REV. 2007 (2010) (proposing a unified conception of privacy rights across tort and other branches of privacy law).

15. *Cf.* RONALD DWORKIN, TAKING RIGHTS SERIOUSLY 116–18 (1978) (explaining that the theory that best fits and justifies an area of law will often not fit all of our judgements about it; rather, to achieve coherence, it will often find some judgements to be mistaken or misconceived).

16. Carpenter v. United States, 138 S. Ct. 2206, 2221–23 (2018).

17. Kyllo v. United States, 533 U.S. 27, 40–41 (2001).

18. Privacy rights are a type of informational right, not vice versa. For this reason, data protection law is not the EU's version of privacy law, as is often suggested. Because data protection law is far more expansive, it remains outside the scope of this Article, and it may not fit within any single unifying theory—though that is a question that must be set aside.

these interests are differentiated, we may find that none justifies protections that are as expansive as those that have been envisioned. Thus, clarifying the nature of privacy losses and violations, and the difference between them, reveals not only the unity but also the limits of privacy.

II. A TAXONOMY OF PRIVACY SCHOLARSHIP

A review of the vast body of privacy scholarship might lead one to conclude that it contains intractable conflict about the nature of privacy,[19] but there is actually less disagreement than at first appears. Much of the apparent conflict arises from a failure to differentiate between normative and descriptive accounts of privacy, the variations within them, and the different questions to which they provide answers. A taxonomic analysis of these differences (in this Part) helps reveal the errors that conflation causes (in the next Part). While some of the individual themes identified in this survey of the literature have been noted by others, many of the most important distinctions have not.[20]

A. NORMATIVE ACCOUNTS

Most of the privacy literature is devoted to the development of normative accounts of privacy (as a matter of law or morality) that address one or more of the following: the interests that privacy protects, the rights that arise from these interests, and the domain of these rights.

1. The Interests that Privacy Protects

A significant body of privacy scholarship has defined privacy in terms of the interests it protects. Three sets of interconnected interests—individual, relational, and societal—have received significant attention.[21]

An individual's interest in being an autonomous person is at the core of many definitions of privacy. For example, privacy has been defined as protecting "inviolate personality,"[22] "the individual's interest in becoming, being, and remaining a person,"[23] "the individual's independence, dignity and integrity,"[24] and other similar values.[25]

19. *See supra* notes 2–3; *see infra* Section II.A.

20. *Cf.* Solove, *Conceptualizing Privacy*, *supra* note 3, at 1099–124 (differentiating between conceptions of privacy based in: the right to be let alone, limited access to the self, secrecy of personal information, control over personal information, personhood, and intimacy).

21. *See generally* NISSENBAUM, *supra* note 13, at 74–77, 84–88 (identifying these three categories).

22. Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 205 (1890).

23. Jeffrey H. Reiman, *Privacy, Intimacy, and Personhood*, 6 PHIL. & PUB. AFF. 26, 44 (1976).

24. Edward J. Bloustein, *Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser*, 39 N.Y.U. L. REV. 962, 971 (1964).

25. *See, e.g.*, Joseph Kupfer, *Privacy, Autonomy, and Self-Concept*, 24 AM. PHIL. Q. 81, 81 (1987) ("development of an autonomous self"); Parent, *supra* note 6, at 278 ("freedom and individuality"); Jed Rubenfeld, *The Right of Privacy*, 102 HARV. L. REV. 737, 784 (1989) ("freedom not to have one's life too totally determined by a progressively more normalizing state").

It is widely agreed, however, that privacy does not just protect the individual in isolation, but also the personal relationships that are essential to human flourishing. Often, these interests are identified in terms of intimacy.[26] For example, Tom Gerety defines privacy as "control over the intimacies of personal identity"[27] and Julie Inness identifies intimacy as the defining feature unifying the set of intrusions that are properly called privacy invasions.[28] But others characterize the relational interests more broadly. For example, Charles Fried defines privacy as protecting relationships of "respect, love, friendship and trust."[29]

Finally, extending beyond personal relational interests, there are the interests of the individual in participating in social and political life. For example, Julie Cohen suggests that privacy fosters a capacity for autonomy that "is an indispensable condition for reasoned participation in the governance of the community and its constituent institutions."[30] Highlighting that "freedom from surveillance . . . is foundational to the practice of informed and reflective citizenship," she argues that privacy "is an indispensable structural feature of liberal democratic political systems."[31] Others have also developed this general position.[32]

2. The Rights that Arise from Privacy Interests

Along with providing an account of the interests that privacy protects, a normative account of privacy must identify the moral or legal rights we have by virtue of these interests.[33] This is perhaps the most contentious issue in the privacy literature, with four broadly different positions being advanced.

---

26. *See* Solove, *Conceptualizing Privacy*, *supra* note 3, at 1121 (providing an overview of intimacy-based theories of privacy).

27. Gerety, *supra* note 2, at 236.

28. INNESS, *supra* note 10, at 56; *see also* JEFFREY ROSEN, THE UNWANTED GAZE: THE DESTRUCTION OF PRIVACY IN AMERICA 7–9 (2000) (explaining that intimate information is meant to stay within the context of close relationships); Robert S. Gerstein, *Intimacy and Privacy*, 89 ETHICS 76, 81 (1978) ("[I]ntimacy simply could not exist unless people had the opportunity for privacy.").

29. Charles Fried, *Privacy*, 77 YALE L.J. 475, 477 (1968); *see also* James Rachels, *Why Privacy Is Important*, 4 PHIL. & PUB. AFF. 323, 326 (1975) (arguing that privacy is needed "to create and maintain different sorts of social relationships with different people").

30. Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 STAN. L. REV. 1373, 1426 (2000).

31. Julie E. Cohen, *What Privacy Is For*, 126 HARV. L. REV. 1904, 1905 (2013).

32. *E.g.*, PRISCILLA M. REGAN, LEGISLATING PRIVACY: TECHNOLOGY, SOCIAL VALUES, AND PUBLIC POLICY 213 (1995) ("Privacy . . . has value not just to the individual as an individual or to all individuals in common but also to the democratic political system."); SOLOVE, *supra* note 2, at 93; Gavison, *supra* note 10, at 423 (arguing that privacy is important because it promotes "liberty, autonomy, selfhood, and human relations, and furthering the existence of a free society").

33. Another underlying question here is whether and how privacy rights are distinct from other rights that seem to provide similar protections. *Compare* Judith Jarvis Thomson, *The Right to Privacy*, 4 PHIL. & PUB. AFF. 295, 310 (1975) (arguing that privacy rights are reducible to other

The first position builds on the foundational claim of Warren and Brandeis that the right of privacy is the right "to be let alone."[34] This conception of the right of privacy, which treats it as a negative right, is fairly abstract; but it has been developed to include concrete rights restricting privacy intrusions and the disclosure of private information.[35] These rights have been recognized to varying degrees in common law, constitutional law, and statute.[36] Recently, Lior Strahilevitz has proposed combining these rights into a single right that could be used as the basis for reunifying privacy law across multiple areas.[37]

The second position, which develops privacy into a positive rather than negative right, defines it as a right to informational control.[38] This conception of privacy became influential in the 1960s through the work of Alan Westin, who argued that "[p]rivacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others,"[39] as well as Charles Fried, who defined privacy as "the *control* we have over information about ourselves."[40] In the following decades, this control-based definition was widely advanced,[41] sometimes with further refinements.[42]

The third position characterizes informational control as just one aspect of a broader privacy right—a right that encompasses various forms of decisional autonomy. For example, Julie Inness defines privacy as "the state of possessing control over a realm of intimate decisions, which includes decisions about intimate access, intimate information, and intimate

rights), *with* Thomas Scanlon, *Thomson on Privacy*, 4 PHIL. & PUB. AFF. 315, 315 (1975) (rejecting Thomson's reductionism).

34. Warren & Brandeis, *supra* note 22, at 205.

35. These are the two core privacy rights identified by Prosser in his foundational work. William L. Prosser, *Privacy*, 48 CALIF. L. REV. 383, 389 (1960). I set aside his other two privacy torts (misappropriation of identify and false light), as Prosser admits that they do not cohere, and his suggestion that they are privacy rights is often rejected. *See, e.g.*, RICHARD A. POSNER, THE ECONOMICS OF JUSTICE 272–73 (1981) (identifying seclusion and concealment as the core of privacy); Gerety, *supra* note 2, at 246–81 (arguing that only the intrusion and disclosure torts are truly concerned with privacy).

36. *See* Strahilevitz, *supra* note 14, at 2015–24.

37. *See id.* at 2010–11.

38. *See generally* Solove, *Conceptualizing Privacy*, *supra* note 3, at 1109–15 (discussing this approach).

39. ALAN F. WESTIN, PRIVACY AND FREEDOM 7 (1967).

40. Fried, *supra* note 29, at 482.

41. *See, e.g.*, ARTHUR R. MILLER, THE ASSAULT ON PRIVACY: COMPUTERS, DATA BANKS, AND DOSSIERS 25 (1971) (defining privacy as "the individual's ability to control the circulation of information relating to him").

42. For example, Richard Parker specified: "The definition of privacy defended in this article is that privacy is control over when and by whom the various parts of us can be sensed by others. By 'sensed,' is meant simply seen, heard, touched, smelled, or tasted." Richard B. Parker, *A Definition of Privacy*, 27 RUTGERS L. REV. 275, 281 (1974) (emphasis omitted).

actions."[43] Likewise, the Supreme Court originally advanced this conception of privacy in its decisions recognizing constitutionally-protected rights to contraception and abortion, which it characterized as privacy rights.[44] However, it no longer does so,[45] in line with the widespread criticism that the interest at stake in these cases is autonomy, not privacy.[46]

The fourth approach to privacy rights argues that they are context-dependent rights governing the transmission of personal information. For example, Helen Nissenbaum argues "that a right to privacy is neither a right to secrecy nor a right to control but a right to *appropriate* flow of personal information,"[47] and that the content of this right is defined in terms of social norms: "Inappropriate information flows are those that violate context specific informational norms . . . a subclass of general norms governing respective social contexts."[48] These norms differ depending on the type of information at issue, the actors involved, and the principles under which the information is transmitted.[49] Highlighting the role of context in determining the content of privacy rights, Nissenbaum suggests that the right to privacy is a right to "contextual integrity."[50]

3. The Domain of Privacy Rights

In addition to providing an account of the content of privacy rights, a normative account of privacy might specify, and thereby limit, the conduct or matters to which they apply—what might be called the domain of privacy rights. For example, it is often suggested that privacy rights only apply to certain types of information that can be properly classified as private.[51] In specifying what information counts, many privacy scholars have identified "intimacy" as a key criterion.[52] But this approach has been criticized for failing

43. INNESS, *supra* note 10, at 140.

44. *See* Roe v. Wade, 410 U.S. 113, 152–55 (1973); Griswold v. Connecticut, 381 U.S. 479, 485 (1965).

45. Jamal Greene, *The So-Called Right to Privacy*, 43 U.C. DAVIS L. REV. 715, 717–18 (2010).

46. *See, e.g.*, TRIBE, *supra* note 11, at 1352; Gavison, *supra* note 10, at 438–39; Gross, *supra* note 6, at 38; Henkin, *supra* note 11, at 1410–11.

47. NISSENBAUM, *supra* note 13, at 127.

48. *Id.* at 9.

49. Helen Nissenbaum, *Respecting Context to Protect Privacy: Why Meaning Matters*, 24 SCI. & ENGINEERING ETHICS 831, 839 (2018).

50. *Id.* at 839–40. A related position is advanced by Solove, who argues: "Privacy is a dimension of certain practices and aspects of life. . . . Privacy invasions disrupt and sometimes completely annihilate certain practices." Solove, *Conceptualizing Privacy*, *supra* note 3, at 1129. *See generally* Robert C. Post, *The Social Foundations of Privacy: Community and Self in the Common Law Tort*, 77 CALIF. L. REV. 957 (1989) (defining privacy rights in terms of social norms of civility).

51. *See, e.g.*, Solove, *Conceptualizing Privacy*, *supra* note 3, at 1104. A related normative claim is that information can be private even if it is known to others. *Id.* at 1108–09.

52. *See, e.g.*, INNESS, *supra* note 10, at 56 ("[P]rivacy is the state of the agent having control over a realm of intimacy, which contains her decisions about intimate access to herself (including intimate informational access) and her decisions about her own intimate actions.").

to protect other important aspects of privacy.[53] There is also a growing view that the domain of privacy rights cannot be defined categorially, but rather must be responsive to context.[54]

B. DESCRIPTIVE ACCOUNTS

While recent debates about privacy have focused on the normative questions identified above, privacy is not only a normative concept of law and morality, but also a descriptive one. It refers not only to rights and interests that can be violated, but also to a state of affairs (or, a "condition") that can be lost. In having two senses—one normative, one descriptive—privacy is similar to many other important moral and legal concepts, such as liberty.[55] This distinction is now often overlooked, but it was once widely recognized as important.[56]

Those who have identified and explored the descriptive side of privacy have generally defined it in terms of "limited access" to some dimension of one's self. For example, Hyman Gross defined the condition of privacy as "the condition of human life in which acquaintance with a person or with affairs of his life which are personal to him is limited,"[57] David O'Brien defined it as "an existential condition of limited access to an individual's life experiences and engagements,"[58] and William Parent defined it as "the condition of not having undocumented personal knowledge about one possessed by others."[59]

53. *See, e.g.*, Gerety, *supra* note 2, at 281 n.175 (arguing that it is a mistake to think that privacy is "reducible to what may be its paradigm, sexual intimacy").

54. *See, e.g.*, Parker, *supra* note 42, at 279 ("It is tempting to try and limit the definition of privacy to control over certain items of information. But this approach is a mistake. Although there is some information which seems peculiarly related to privacy . . . , a loss of control over most items of information about ourselves is sometimes related to privacy and sometimes not."); *see also* Marc Jonathan Blitz, *Video Surveillance and the Constitution of Public Space: Fitting the Fourth Amendment to a World that Tracks Image and Identity*, 82 TEX. L. REV. 1349, 1412 (2004) (providing examples).

55. *See supra* note 5 and accompanying text.

56. *See, e.g.*, Gross, *supra* note 6, at 35 ("Privacy is a state of affairs, and before we speak of 'rights of' or 'interests in' or 'invasions of' it, we ought to be acquainted with its distinguishing features."); O'Brien, *supra* note 3, at 69, 74–75 (criticizing the privacy literature's "failure to adequately distinguish between the concept of privacy and a right to privacy"); W.A. Parent, *A New Definition of Privacy for the Law*, 2 LAW & PHIL. 305, 309 (1983) ("The concept of a right to privacy is quite different from and should not be confused with the concept of privacy simpliciter."); Schoeman, *supra* note 6, at 3 (highlighting the importance of differentiating "the question of whether or not one has undergone a loss of *privacy* from the question of whether or not one's *right to privacy* has been infringed or violated"). The distinction is now often overlooked, or if acknowledged, dismissed as unimportant. *See, e.g.*, Moore, *supra* note 6, at 421 (concluding that "we should not be overly worried about defining a state or condition precisely"). But there are exceptions. *See, e.g.*, Powers, *supra* note 6, at 372 (arguing for a definition that "seeks to identify a common analytic basis for all privacy claims").

57. Gross, *supra* note 6, at 35–36 (emphasis omitted).

58. O'Brien, *supra* note 3, at 75.

59. Parent, *supra* note 6, at 269.

As I will ultimately advance a position that falls within this general category, I will at this point merely highlight three issues in relation to which limited access accounts have diverged. First, there is the question of whether access can be completely defined in informational terms, or whether non-informational access can cause privacy losses.[60] Second, there is the matter of whether a person must actually be accessed—informationally or otherwise—for a privacy loss to occur, or whether mere accessibility is sufficient.[61] Third, there is the issue of whether the limitation on access or accessibility must be desired by the person to count as privacy, or whether privacy can be imposed.[62]

It is also possible to depart entirely from the limited access framework in defining the condition of privacy. For example, drawing on normative conceptions of privacy, one could argue that "informational control" describes not only the right of privacy, but also the condition of privacy. On this account, privacy would not exist, as a descriptive matter, when an individual lacks informational control. However, this approach is uncommon—and for good reasons, which I will discuss in Part IV, after first demonstrating the significant errors caused by conflating privacy losses and violations.

III. CONFLATION ERRORS

Having identified the distinction between normative and descriptive accounts of privacy—and the different questions they might seek to answer—I will now show how their conflation has given rise to two related errors in a significant body of privacy literature: mistargeted critique, in which descriptive accounts are rejected for failing to answer normative questions; and misguided skepticism, in which the failure to differentiate between them

60. For example, Ruth Gavison rejects the pure informational approach, arguing:

> A loss of privacy occurs as others obtain information about an individual, pay attention to him, or gain access to him. These three elements of secrecy, anonymity, and solitude are distinct and independent, but interrelated, and the complex concept of privacy is richer than any definition centered around only one of them.

Gavison, *supra* note 10, at 428–29.

61. For example, Anita Allen tracks Gavison's three-prong definition, except for the fact that she adopts a criterion of limited inaccessibility rather than limited access: "My own restricted-access definition of 'privacy' is this: personal privacy is a condition of inaccessibility of the person, his or her mental states, or information about the person to the senses or surveillance devices of others." ALLEN, *supra* note 10, at 15. The difference between access and accessibility is highly significant for reasons discussed in Section IV.A.

62. For example, Sissela Bok tracks Gavison's three-prong definition, but with the additional requirement that the lack of access across the three dimensions be *desired* by the person at issue: "I shall define privacy as the condition of being protected from unwanted access by others—either physical access, personal information, or attention." SISSELA BOK, SECRETS: ON THE ETHICS OF CONCEALMENT AND REVELATION 10–11 (1983). *But see* ALLEN, *supra* note 10, at 27 ("Privacy aptly describes even some conditions of unwanted inaccessibility.").

has resulted in the mistaken conclusion that a unified theory of privacy rights is unattainable.

A. MISTARGETED CRITIQUE

A large body of critical scholarship suffers from a type of logical error known as a "category mistake," in which something that belongs in one category is treated as though it belongs in a different category.[63] Often, category mistakes occur when an object of critique is treated as though it has, or should have, a property that it cannot have. In the privacy literature, this is widespread. Descriptive accounts of privacy are often rejected for failing to provide answers to normative questions. This can be seen in three common critiques of limited access accounts of privacy, which track the taxonomic analysis developed above.

In the first critique, limited access accounts are rejected for failing to explain the interests that privacy protects. For example, Dan Solove argues that this approach should generally be rejected on the grounds that it fails to explain the "value of privacy" and is therefore unable to answer important questions about the nature of privacy rights and private matters.[64] Others make similar claims.[65] It should now be clear, however, that these complaints are misguided. They fail to recognize that limited access accounts of privacy are descriptive, not normative, accounts.

In the second critique, the limited access approach is rejected for failing to provide an account of privacy rights in various ways. For example, it has been rejected for failing to provide criteria by which to differentiate between legitimate and illegitimate modes of acquiring information,[66] identify clear cases of privacy violations,[67] and explain what is important about privacy

63. *See generally* Ofra Magidor, *Category Mistakes*, *in* THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., Fall 2019 ed.), *available at* https://plato.stanford.edu/archives/fall2019/entries/category-mistakes [https://perma.cc/AT9W-SLMF] (explaining category mistakes).

64. Solove, *Conceptualizing Privacy*, *supra* note 3, at 1104; *see* SOLOVE, *supra* note 2, at 20.

65. *See, e.g.*, INNESS, *supra* note 10, at 44–45 (arguing that limited access theories should be rejected on the grounds that they define privacy in a value-neutral way); Judith Wagner DeCew, *The Scope of Privacy in Law and Ethics*, 5 LAW & PHIL. 145, 152 (1986) (rejecting William Parent's limited access definition because it provides "no way . . . to judge what should or should not be a part of the public record" and "leaves no room for a normative sense of privacy encompassing interests worthy of protection" (emphasis omitted)).

66. *See, e.g.*, DeCew, *supra* note 65, at 146–54 (rejecting Parent's definition on these grounds); *see also* SOLOVE, *supra* note 2, at 20 (arguing that a problem with limited access theories is that they provide "no understanding as to the degree of access necessary to constitute a privacy violation").

67. *See, e.g.*, INNESS, *supra* note 10, at 46–47 (arguing that limited access theories should be rejected because they suggest that a person's privacy is not violated when they must hide to avoid being seen by a Peeping Tom). Likewise, when Inness rejects limited access theories on the grounds that privacy and access are not opposed if privacy's function is "to provide the individual

rights.[68] But again, these criticisms are based on a category mistake. They expect an account of the condition of privacy to be an account of the right of privacy.

In the third critique, limited access accounts are rejected for failing to identify the domain of privacy rights. There are two general versions of this critique. The first suggests that limited access accounts are under-inclusive because personal facts can be known to others, but private.[69] The second version of the critique suggests that they are over-inclusive because personal facts can be unknown to others, but not private.[70] Setting aside the validity of the assumption that some *types* of information can be categorically classified as private (and others as not), which there is good reason to reject,[71] the problem I want to highlight is that these critiques assume a normative conception of privacy. What constitutes a "private matter" is determined by reference to a conception of the value of privacy or an understanding of the types of information people often want to keep private (i.e., in terms of interests in privacy). Thus, it is a category mistake to suggest that these critiques identify a problem with limited access accounts of privacy, which are descriptive and intentionally set aside the question of when privacy should be protected.

In short, limited access accounts are often rejected for failing to explain the interests that privacy protects, the rights that arise from these interests, and the domain of these rights—or in other words, for failing to answer the

with control over certain aspects of her life," she makes a claim about privacy rights (and their function), not the condition of privacy. *Id.* at 6.

68. *See, e.g.*, Steve Matthews, *Privacy, Separation, and Control*, 91 MONIST 130, 141–42 (2008) ("When we say . . . it is important to respect a person's privacy, we surely do not mean it is important to respect the mere condition someone is in of being secluded from us. . . . What we are respecting is the person's explicitly expressed choice, . . . or a choice we must presume they would reasonably make . . . .").

69. *See, e.g.*, DeCew, *supra* note 65, at 155 ("[P]rivate information about one's debts or odd behavior may be publicized. Although it is no longer concealed, it is no less private."); Solove, *Conceptualizing Privacy*, *supra* note 3, at 1109 ("The books we read, the products we buy, the people we associate with—these are often not viewed as secrets, but we nonetheless view them as private matters."). Solove also attributes this position to Stanley Benn, though Benn in fact differentiates between the descriptive issue of whether something is done "in private" (which he defines in terms of informational access) and the normative question of whether something is a "privacy matter" (which he states is both norm-dependent and norm-invoking). Stanley I. Benn, *Privacy, Freedom, and Respect for Persons*, *in* PHILOSOPHICAL DIMENSIONS OF PRIVACY: AN ANTHOLOGY, *supra* note 6, at 223–24.

70. *See, e.g.*, DeCew, *supra* note 65, at 155 ("[W]hatever is secret is concealed or withheld from others, and it may not always be private.").

71. The problem is that there is not any type of information that is categorically related, or unrelated, to privacy. While there are some types of personal information that most people want to keep private (for example, information about sexual habits), this speaks to the value of privacy. Further, this value is ultimately determined by the context of a factual disclosure—not the type of fact disclosed. *See supra* note 55 and accompanying text.

three questions addressed by normative accounts identified above.[72] In these ways, the common critiques all entail category mistakes.

A related way in which these critiques are unjustified, which deserves brief mention, is that they fault limited access theorists for failing to answer normative questions that they do in fact answer. For example, Judith DeCew suggests that William Parent's account provides "no way . . . to judge what should or should not be a part of the public record,"[73] but Parent does address the "criteria of wrongful invasion" in his discussion of the right to privacy.[74] Likewise, Parent and others, such as Ruth Gavison and Hyman Gross, devote significant attention to the value of privacy, after they have first defined the condition of privacy in value-neutral terms.[75] They also clearly explain that this will be the structure of their analyses. For example, Gross explains: "Privacy is a state of affairs, and before we speak of . . . 'interests in' or 'invasions of' it, we ought to be acquainted with its distinguishing features."[76]

Thus far, I have argued that the standard critiques of the dominant descriptive accounts of privacy are mistargeted. In response, one might argue that the entire project of developing a descriptive understanding of privacy does not make sense—and that this is what underlies the critiques I have identified.[77] But there would be two problems with this response. First, the critiques that have been offered do not actually address the question of whether it makes sense to have a non-normative account of privacy.[78] Second, there are good reasons to maintain the distinction including: the term privacy is used in both ways in ordinary language and academic writing; the failure to recognize this has produced significant confusion in the literature; and the distinction allows one to separate questions that should be capable of having different answers. At the very least, a definition of privacy should allow for

72. *See supra* Section II.A.

73. DeCew, *supra* note 65, at 152.

74. Parent, *supra* note 6, at 278–88. Likewise, DeCew argues that Parent's "descriptive emphasis . . . leaves no room for a normative sense of privacy encompassing interests worthy of protection." DeCew, *supra* note 65, at 152 (emphasis omitted). But again, Parent does in fact address this issue when discussing the value of privacy and the right to privacy. Parent, *supra* note 6, at 278–88.

75. Gavison, *supra* note 10, at 440–55; Parent, *supra* note 6, at 275–77.

76. Gross, *supra* note 6, at 35.

77. *Cf.* INNESS, *supra* note 10, at 44 (arguing that value-neutrality is a reason to reject limited access theories of privacy); Moore, *supra* note 6, at 414 (stating that descriptive accounts are "largely uninteresting").

78. Moore briefly touches on this issue in discussing Parent's non-normative definition. Moore, *supra* note 6, at 416. However, he primarily assumes a normative conception is necessary and criticizes Parent for failing to provide one. *Id.* He acknowledges that it "may also be helpful to define a condition of privacy" but concludes that if his "critique of Parent's non-normative conception of privacy is compelling, then we should not be overly worried about defining a state or condition precisely." *Id.* at 421. Further, he does not address any of the arguments made by Parent or others about why it is important to differentiate and develop descriptive and normative theories of privacy. *See supra* note 57.

different answers to the questions of whether privacy *is* protected in a given situation and whether it *should* be protected. The definition should not entail the view that privacy should always be protected.[79] Furthermore, as the next Section will demonstrate, an understanding of the condition of privacy can provide the foundation for a unified theory of privacy rights.

B. MISGUIDED SKEPTICISM

1. Denying Coherence

The failure to distinguish between normative and descriptive conceptions of privacy has generated not only mistargeted critique, but also misguided skepticism about the possibility of developing a unified theory of privacy. This skepticism has taken two general forms. Both conclude that we should not seek a unified conception of privacy or privacy rights, but they reach this conclusion for different reasons.

The first form of skepticism rejects the coherence of privacy on the grounds that there is nothing distinctive about the concept of privacy —nothing that makes it unique from other concepts. This position, sometimes called "reductionism," but more accurately called "eliminativism,"[80] suggests that "our concept of privacy highlights different and unrelated interests in the various contexts in which it applies," such that "all talk of privacy could (and perhaps should) be eliminated in favor of talk of the unrelated interests."[81] For example, Judith Thomson argues that "every right in the right to privacy cluster is also in some other right cluster" and that "the wrongness of every violation of the right to privacy can be explained without ever once mentioning it."[82] For this reason, Thomson concludes, "there is no need to find the that-which-is-in-common to all rights in the right to privacy cluster and no need to settle disputes about its boundaries."[83]

79. Going even further, Parker suggests that an adequate definition of privacy must allow us to differentiate between

> five questions: (1) whether a person *has* lost or gained privacy, (2) whether he *should* lose or gain privacy, (3) whether he *knows* that he has lost or gained privacy, (4) whether he *approves or disapproves* of the loss or gain, and (5) how he *experiences* that loss or gain.

Parker, *supra* note 42, at 278.

80. The contrary view, non-eliminativism, can take two forms: (1) fundamentalism, which maintains that the interest protected by privacy is an irreducible and sui generis interest; and (2) reductionism, which maintains that the interest protected by privacy can be explained in terms of other interests, but that the concept of privacy cannot be eliminated in favour of a more fundamental concept. David Matheson, *A Distributive Reductionism About the Right to Privacy*, 91 MONIST 108, 108–09 (2008); *see also* Powers, *supra* note 6, at 372 (explaining fundamentalism and reductionism).

81. Matheson, *supra* note 80, at 108.

82. Thomson, *supra* note 33, at 312–13.

83. *Id.*; *see also* Anderson, *supra* note 4, at 82; Richard Volkman, *Privacy as Life, Liberty, Property*, 5 ETHICS & INFO. TECH. 199, 199 (2003).

The second form of skepticism rejects the eliminativist position and maintains that privacy is distinctive, but nevertheless denies its coherence. Dan Solove has best developed this position, arguing that most privacy theorists have created confusion by adopting what he calls the "traditional method" of conceptual analysis in which they "attempt to articulate what separates privacy from other things, what makes it unique, and what identifies it in its various manifestations."[84] Solove argues that it is a mistake to search for "the 'essence' of privacy"[85] and the "common set of necessary and sufficient elements that single out privacy as unique."[86] According to Solove, we should instead understand privacy in terms of Ludwig Wittgenstein's notion of "family resemblances," which suggests that there is no common feature or essence of all of the things that we classify with a given concept.[87] Solove argues that privacy theorists should adopt a pragmatic orientation and focus on understanding the features of different types of activities that are said to pose "privacy problems."

While Solove's approach to privacy is illuminating, it is important to recognize that it maintains the distinctiveness of the concept of privacy by denying that it has any core meaning. On his account, an activity counts as a privacy problem, or not, based solely on whether it has "achieved a significant degree of social recognition" as such.[88] Thus, while his taxonomy of privacy problems sheds light on practices that people classify as privacy problems, it is not clear that all of them are properly classified as such—which points to significant cost of adopting Wittgenstein's anti-essentialism in a theory of a legal or moral concept, such as privacy. (A theory of language is another matter).

If one accepts that pure social convention determines the correct use of a legally- or morally-relevant concept, one loses an important means of critiquing societal consensus about how the concept is applied. An example from another context helps illustrate the problem. Imagine that I am confronted with a society in which everyone classifies homosexuality as a disease. If I want to convince them that this classification is incorrect, a standard way to do so would be to point out that all of the other diseases the society recognizes share a set of core characteristics (or satisfy a set of criteria) that homosexuality does not. This line of argument would not be impossible, however, if I took the view that concepts are not defined by core characteristics/criteria that determine the correct scope of their use. I might

84. Solove, *Conceptualizing Privacy*, *supra* note 3, at 1095.

85. *Id.* at 1096.

86. *Id.* at 1095.

87. *Id.* at 1096–97.

88. SOLOVE, *supra* note 2, at 101–02; *see also id.* at 172 (acknowledging this aspect of his account); M. Ryan Calo, *The Boundaries of Privacy Harm*, 86 IND. L.J. 1131, 1141 (2011) (providing a valuable critique of this aspect of Solove's position).

try to identify other reasons why the society should change its use of language, but it is hard to see how any other type of reason would be equally compelling.

Likewise, in the privacy context, Solove's anti-essentialist account means there can be no solid grounds for arguing that a society's views about matters of privacy are incorrect. If something is not recognized as a privacy issue, one cannot appeal to the core of privacy to argue that it should be; and conversely, if something is recognized as a privacy issue, one cannot appeal to the core to argue that it should not be.[89]

This anti-essentialism also precludes other forms of analysis that are important under a traditional understanding of legal and moral concepts. Take, for example, the question of whether the constitutionally protected right to an abortion is properly classified as a privacy-based right or an autonomy-based right. On Solove's approach, the answer to this would depend solely on social convention, which is a problem given that it is relevant to the grounding of the right in the text of the Constitution.[90] In addition, the classification of a claim as a privacy claim would tell us nothing about whether or why we should respect it. Under a traditional theory of privacy, the fact that a claim is properly classified as a privacy claim will be morally and legally significant; if I think privacy is important, I will have reason to think the claim is important. Under Solove's approach, by contrast, the fact that a claim is classified as a privacy claim would not tell me anything about whether I should treat it as morally or legally significant.

In response to these criticisms, one might argue that although Solove's Wittgensteinian approach has these limitations, they are unavoidable if one wants to adopt a theory that can capture the heterogeneity of privacy that the literature has recognized. This would be incorrect. The heterogeneity does not mean that privacy lacks a distinctive core (as Thomson concludes) or unified core (as Solove concludes). To identify the problem with both of these conclusions, it is necessary to first differentiate between two ways in which a theory of privacy could potentially cohere.[91]

2. Regrounding Coherence

The first possible form of coherence, which the privacy literature has generally assumed is the only possibility, is normative coherence. On this approach, coherence would be achieved by identifying the criteria that uniquely justify all the rights we consider to be privacy rights. This is the goal of what Solove calls the "traditional approach" to conceptualizing privacy, and while this goal may not have been actively chosen (as the alternative of

89. *Cf.* Calo, *supra* note 88, at 1141–42 (developing a similar critique of Solove).

90. *See id.*

91. In developing this distinction, I draw on Madison Powers' very helpful distinction between "justificatory reductionism" and "definitional reductionism." Powers, *supra* note 6, at 384–86.

descriptive coherence has generally not been recognized), it is nevertheless widely stated as the goal of theorizing privacy. For example, Reiman states: "What we are looking for is a fundamental interest . . . which provides a basis for a right to privacy to which all human beings are entitled."[92]

The assumption that normative coherence is the goal of theorizing privacy is not only found in those who adopt the "traditional approach," but also in its critics. The growing skepticism about the possibility of developing a unified account of privacy assumes that normative coherence is the only form of coherence. It is only on the basis of this assumption that the heterogeneity of interests protected by privacy rights—and the intractable disagreement in the privacy literature—answers the question of whether a unified account of privacy can be developed. This assumption is unjustified, however, as normative coherence is not the only way in which the concept of privacy might cohere.

The second possibility, which has been largely overlooked in the literature, is descriptive coherence. Under this approach, coherence would be achieved by identifying the state of affairs that is uniquely described as the condition of privacy. The benefit of this approach is that it can accommodate the heterogeneity of the value of privacy. As Alan Rubel explains in another context:

> There is little reason . . . to think that privacy has a single type of value. Privacy regarding one's voting habits with respect to state actors may be an important political value, whereas privacy regarding one's shopping habits with respect to marketers may be instrumentally valuable. Indeed, many instances of privacy loss are likely of no moral concern.[93]

A descriptive account can accommodate this heterogeneity by locating the coherence of privacy in a state of affairs, rather than the value of this state of affairs. Although a descriptive account needs to be supplemented by a normative account if we are to determine when privacy violations occur, this approach "ensures that we are talking about the same subject when we use the language of privacy rights."[94] Thus, before exploring the normative concept of privacy violation, it is important to clarify the descriptive concept of privacy loss.

92. Reiman, *supra* note 23, at 38; *see also* Rachels, *supra* note 29, at 323 ("[T]he first element of a theory of privacy should be 'a characterization of the special interest we have in being able to be free from certain kinds of intrusions.'").

93. Alan Rubel, *The Particularized Judgment Account of Privacy*, 17 RES PUBLICA 275, 287 (2011). Likewise, Madison Powers concludes that we should "doubt that any one value could be adequate to account for all cases in which privacy matters, or to suppose that any grouping of these diverse values uniquely support privacy rights rather than rights of some other sort." Powers, *supra* note 6, at 385.

94. *Id.* at 386.

IV. PRIVACY LOSSES

I have thus far argued that the failure to differentiate between descriptive and normative conceptions of privacy—between privacy as a value-neutral state of affairs that can be lost and privacy as a right that can be violated—has generated misguided critique and skepticism in the literature. In what follows, I will turn to the first core question this analysis raises: the question of how we should understand privacy losses. My analysis of this issue will give special consideration to the epistemological dimensions of privacy. Although privacy has frequently been defined in terms of knowledge,[95] few have explored whether or how the various elements of knowledge are essential to privacy losses (or privacy violations).[96] My argument, in short, will be that a privacy loss occurs when true information about a person is accessed by another via a means that has epistemic merit.[97] There are three essential criteria in this account: access, epistemic merit, and truth. When all three are satisfied, a person experiences a privacy loss in the given piece of information and with respect to the person who accessed it.

A. ACCESS

While it is uncontroversial to state that many core cases of privacy losses arise when one's personal information is accessed by another, my claim that this is a defining criterion faces three possible challenges.

1. The Accessibility Objection

The first challenge comes from theorists who argue that privacy losses can be caused by mere accessibility—that actual access is not required. Anita Allen, for example, argues that privacy is the "condition of inaccessibility of the person, his or her mental states, or information about the person to the

95. This approach dates back at least as far as 1890 when, shortly before Brandeis and Warren published their canonical article, E.L. Godkin defined "the right to privacy" as a person's "right to decide how much knowledge of this personal thought and feeling, and how much knowledge . . . of his own private doings and affairs . . . the public at large shall have." E.L. Godkin, *The Rights of the Citizen.: IV.—To His Own Reputation*, 8 SCRIBNER'S MAG. 58, 65–67 (1890). More recently, knowledge has been incorporated into limited-access definitions of privacy, for example, Gavison, *supra* note 10, at 423 (privacy concerns "the extent to which we are *known* to others" (emphasis added)), as well as control-based definition, for example, Fried, *supra* note 29, at 483 ("Privacy . . . is control over *knowledge* about oneself." (emphasis added)).

96. The one notable exception is a small set of articles published in a special issue of the journal *Episteme* on "Privacy, Secrecy, and Epistemology" published in Volume 10(2) 2013. In addition, many privacy scholars have indirectly made arguments that are relevant to the epistemology of privacy, which I will draw on and refine in my analysis.

97. While a definition of privacy losses might also limit the *types* of facts that count, attempts to identify categories of "private" information generally face problems that have been identified in the literature. *See supra* note 55. However, there might be some minor limits on the type of personal facts that can give rise to privacy losses. For example, as Don Fallis discusses, it seems that a person cannot "lose privacy about the fact that he is self-identical," as that fact is not specific to any person. Don Fallis, *Privacy and Lack of Knowledge*, 10 EPISTEME 153, 156 (2013).

senses or surveillance devices of others."[98] On this account, a privacy loss occurs whenever information becomes accessible to others, regardless of whether the information is ultimately accessed, and cases "of complete though unexploited accessibility are not adequately described as conditions of privacy."[99] If this is right, various new data-mining and processing technologies could generate privacy losses merely by increasing the accessibility of personal data.[100]

Advocates of this position support it with various examples. Imagine, for instance, that a government sets up an extensive camera surveillance system but does not turn it on; an ocean wave pulls off one's bathing suit but no one sees; a stranger finds one's lost diary in a park but does not read it; or a hacker gains access to one's web-browsing history but chooses not to view it.[101] Such cases might seem to demonstrate that privacy can be lost through mere accessibility, as their intuitive pull is clear. But it is a mistake to identify these as cases of privacy losses.

The problem with this position is that it conflates the *condition of privacy* with the *conditions that protect* privacy. These examples of accessibility are clearly cases in which people's privacy is not well protected, and this lack of protection may be relevant to the question of whether any of them can reasonably expect to have privacy in the future. However, even if none of them has a "reasonable expectation of privacy," this is irrelevant to the question of whether they have privacy in the present.

In conflating the condition of privacy with the conditions that protect privacy, the accessibility theory collapses an important qualitative distinction between access and risk of access. Imagine, for example, that you buy a telephoto lens to look at wildlife. The lens could also be used to see otherwise-inaccessible things inside my house, but you do not use it in this way. If the accessibility theory were correct, your mere purchase of the lens would cause me to suffer a privacy loss, as it would increase the accessibility of the information inside my house. In addition, there would be no qualitative difference between this state of affairs and one in which you buy and use the lens for the purpose of looking inside my house. According to the accessibility theory, I would suffer the same *type* of loss in both cases; the only difference would be the magnitude of the loss. It is hard to imagine, however, that anyone would see things this way. Even assuming for the sake of argument

98. ALLEN, *supra* note 10, at 15.

99. *Id.* at 29 (arguing that privacy would be lost, though not completely).

100. A similar conclusion would also follow from defining privacy in terms of a low *likelihood* of access, rather than limited *accessibility*. While the two criteria will often result in the same conclusion—and face similar objections—they would result in different outcomes in some cases. For example, an email left open on a public computer might be easily accessible, but unlikely to be read (e.g., if the computer is rarely used); whereas an email on a secure server might be fairly inaccessible, but likely to be read (e.g., if it contains information desired by hackers).

101. *See, e.g.*, ALLEN, *supra* note 10, at 29; Rubel, *supra* note 93, at 278, 284.

that the purchase of the lens to look at wildlife would cause me to suffer a meaningful loss, it would be a loss in the conditions that limit access to me. This is clearly not the same type of loss I would suffer if you bought and used the lens to look at me.

In response, one might argue that although there is a qualitative difference between the cases I have identified, there is not always one between cases of access and accessibility. Compare, for example, the following cases: (1) person A hacks my email and reads it, and (2) person B hacks my email and is 95 percent likely to read it in the coming day. The cases seem very similar in terms of privacy, and on this basis, one might argue that both access and accessibility can cause privacy losses. But this would be a mistake. For while the cases are very similar, the similarity is not in the dimension of privacy loss, but rather in two other dimensions that are often confused with it: privacy violation and privacy-related harm.

First, the privacy violations are the same in both cases. Both involve the hacking of email, and so entail the violation of the same privacy right. This does not mean, however, that they both entail a privacy loss. Rather, as I will argue in Part V, the best account of the nature of privacy rights defines them by reference to privacy losses, while allowing for the possibility of violations without losses. Thus, the fact that these cases of access and risk of access entail equally wrongful acts does not mean that they both entail privacy losses.

Second, the down-stream harms may be the same in both cases. For example, in both cases, there is a risk that the hackers will publish the emails, generating reputational harms. Further, in both cases, there might be identical risks of this happening. Imagine, for example, that hacker A (who reads the email) is 95 percent likely to publish the details of what he reads, whereas hacker B (who is 95 percent likely to read the email) is 100 percent likely to publish the details of what he reads. In this hypothetical, the risk of the reputational harms would be 95 percent in both cases. This equivalence of the down-stream harms does not, however, mean that access and risk of access both involve privacy losses.

In general, for any given good X (e.g., life, liberty, money, etc.), there are legally and morally recognized distinctions between cases in which the good has been lost versus cases in which it is at risk of being lost. This is not to say that they are distinct in every way,[102] but rather that they differ in some important ways. Compare, for example, a case in which I have lost my life with a case in which I am at high risk of losing it. The former entails a harm that is

102. On the contrary, if both the loss and the risk of loss are caused by wrongful acts, the acts might be treated as equally wrongful by both criminal law and morality; the difference between them might not be relevant. Likewise, the content of our legal and moral duties to prevent losses and risks of losses might be similar or the same. But this does not mean that both cases entail the same type of loss. The question of loss is about the victim rather than the wrongdoer, and from this perspective, there is clearly a qualitative difference between the loss of life and the risk of loss.

recognized by tort law, whereas the latter does not (unless the loss comes to pass).

One might note, however, that in privacy cases (unlike in cases of life and death), risk of access might be worse—from the perspective of the victim—than actual access. Compare, for example, the following two cases: in the first, a nude photo of someone is accessed by one person; in the second, the nude photo is made accessible to 100 people.[103] It seems likely that most people would agree that it would be worse to be in the second situation than the first, and one might argue that this demonstrates that the second entails a privacy loss that is greater than that in the first. But the problem with this conclusion is that it assumes these cases are undesirable in the same way.

The better explanation of these cases is that for any given good X, an actualized small loss can be preferable to a risk of a large loss. This is a basic feature of economic rationality and the assignment of probability-weighted values to outcomes. The probability-weighted value of a small risk of a large loss of X can be greater than the value of a small loss of it. But it does not follow that the risk of a loss of good X, however undesirable, constitutes a loss of *that* good.[104] It might be the case that the risk of the loss of X constitutes a loss of some other good Y (e.g., one's security in X). And for this reason, one might argue that the risk of losing X should itself be recognized as a legally or morally relevant loss. But this is a controversial argument in the risk literature, and even its main proponents do not argue that the risk of losing X amounts to actually losing X.

It seems the only goods that might be exceptions to this principle are those that are themselves defined in terms of risk. Take, for example, the good of "safety." Because safety is defined in terms of risk (i.e., it is defined as a state of affairs in which the risk of harm is low), there is no qualitative difference between "a risk of a loss of safety" and "a loss of safety" (just as there is no qualitative difference between "a risk of a risk of harm" and "a risk of harm"). The only thing that changes when one adds "risk" in front of "loss of safety" is the amount of risk of harm. Thus, to return to an issue that arose in my earlier discussion of the telephoto lens, the question becomes whether the condition of privacy is a condition like safety—whether it is just a state of affairs in which the risk of future harms is reduced. If so, there would be no qualitative difference between access and risk of access.

Upon analysis, however, it becomes clear that this risk-based account of privacy should be rejected for several reasons. The most important is its incompatibility with the widespread view that a loss of privacy (like the loss of

103. This example involves risk (i.e., *ex ante* uncertainty about whether a loss will occur), but the following analysis applies equally to *ex post* uncertainty about whether a loss has occurred.

104. If this were the case and a loss occurred at the time the risk was imposed, it would follow that a gain would occur when the risk did not materialize. But this account would create confusion that is avoided by the normal way of describing this situation; there was a risk of a loss, but the loss did not materialize.

other goods such as liberty, etc.) can be legally and morally significant independently of the risks it creates. For example, I think most would agree that a person can lose something of value when another person sees him naked, even if it is impossible that this will result in any other harm to him. The mere fact of access is considered a meaningful loss. According to the risk-based account, however, access to personal information is only significant to the extent it creates a risk of some other undesirable event, such as reputational harm. Thus, if one thinks access can be undesirable in and of itself—and not merely because it creates a risk of some other undesirable event—then one should reject the view that privacy loss should itself be defined in terms of risk.

The risk-based account of privacy is also incompatible with the widespread use of the concept of privacy to describe the materialization of privacy risks *ex post*, without reference to what might happen in the future. In this sense, privacy is unlike the concept of safety, which can only be applied to the world *ex ante*. It is meaningful to talk about a lack of privacy even if there is no risk of a further harm, whereas it is not meaningful to talk about a lack of safety in this way. A comparison best illustrates this difference: when an unsafe building collapses and hurts someone who is then at no further risk of injury, we do not say the person lacks safety in this regard; but when an embarrassing piece of information is disclosed to the world and all possible harm is done, we do say the person lacks privacy in this regard.

Further, this fundamental difference between safety and privacy does not only arise in cases in which there is no longer any risk of further harm. Compare, for example, a privacy case in which a person's email is read following the materialization of a privacy risk with a safety case in which a person is burned by acid following the materialization of a safety risk. In both cases, it is possible that the materialization of the risk creates the risk of further harms (e.g., reputational harms in the email case and health harms in the acid case). But when referring to the event that has occurred, from an *ex post* perspective, we would say that the first person has lost privacy with respect to the information in the email, but not that the second person has lost safety with respect to the acid. Thus, the concept of privacy loss is not fundamentally risk-based like safety loss.

For all of the above reasons, privacy losses should not be defined in terms of accessibility. This account does not provide a viable challenge to my claim that privacy losses turn on actual access.

2. The Control Objection

The second challenge to my claim that access is necessary for privacy losses comes from privacy theorists who argue that privacy consists of control over access to oneself. While this account of the nature of privacy has generally been advanced in normative accounts of privacy (and is best interpreted as an

argument about the content of privacy rights),[105] it has also appeared in some accounts of the condition of privacy.[106] If this account were right, it would pose a challenge to my claim that privacy losses turn on access. But as I will argue in this Section, this use of a control-based account (to describe the condition of privacy, rather than the content of privacy rights) faces significant problems.

One problem is that control is clearly not sufficient for privacy.[107] Consider, for instance, a case in which a person intentionally shares previously secret information about herself with a group of people who are guaranteed to seek her permission before using or sharing the information with anyone else. In this scenario, there would be no change in her control over her information. But it seems uncontroversial to say that in this case, she has experienced a small loss of privacy (though not a privacy violation, as she chose to share the information). A fact that was completely private is no longer private with respect to a small number of people with whom she has shared it (though it remains private with respect to the rest of the world). This is a case of a small, controlled privacy loss. Thus, privacy cannot be defined simply as control over information about oneself.

To avoid this problem, the requirement of control might be interpreted as supplementing, rather than replacing, the requirement of limited access. In support of this interpretation, one might rely on the work of Charles Fried, who argues:

> As a first approximation, privacy seems to be related to secrecy, to limiting the knowledge of others about oneself. This notion must be refined. It is not true, for instance, that the less that is known about us the more privacy we have. Privacy is *not simply* an absence of information about us in the minds of others; rather it is the control we have over information about ourselves.[108]

While Fried's claim here is best interpreted as referring to the right to privacy, advocates of a control-based definition of the condition of privacy might make a similar claim.

105. Many have made this observation. *See, e.g.*, Moore, *supra* note 6, at 417 ("I do not think that control-based privacy theorists actually intend to support a purely non-normative conception of privacy."); O'Brien, *supra* note 3, at 74 (arguing that a control-based definition of privacy confuses the condition and right to privacy); Parent, *supra* note 6, at 273 n.11 ("Proponents of a control definition might respond by saying that they are really interested in identifying *the right to privacy* . . . . But then they should have said so explicitly instead of formulating their contention in terms of privacy alone.").

106. For example, Louis Lusky states: "Privacy is the condition enjoyed by one who can control the communication of information about himself." Louis Lusky, *Invasion of Privacy: A Clarification of Concepts*, 72 COLUM. L. REV. 693, 709 (1972).

107. *See, e.g.*, O'Brien, *supra* note 3, at 74; Parent, *supra* note 6, at 273.

108. Fried, *supra* note 29, at 482 (emphasis added and omitted).

In support of adding a requirement of control to the requirement of limited access, one might cite examples that are similar to those used in support of the limited accessibility theories discussed above. For instance, Alan Rubel writes: "[I]t is difficult to maintain that one has privacy regarding one's naked body on a nude beach, even if others fail to see it. The control view explains why: one lacks privacy because others noticing or not is beyond one's control."[109] The problem with this type of example, however, is that it involves both loss of control and an increased risk of disclosure. Thus, people's intuitions about it might be driven by either an accessibility theory of privacy loss, which should be rejected for the reasons identified above,[110] or a control theory.

To avoid this problem, the proponent of a control criterion needs to offer an example in which there is neither access nor accessibility, but in which a person lacks privacy because he lacks control. This would indicate that, at least in some cases, control is necessary for privacy. Such an example is, however, difficult to find in the literature. Perhaps the most promising possibility comes from Fried's explanation of why control is a necessary element of privacy, which he supports with the example of a man on a desert island: "To refer . . . to the privacy of a lonely man on a desert island would be to engage in irony. The person who enjoys privacy is able to grant or deny access to others."[111]

While Fried may not have offered this desert island example to support the claim that control is necessary for the *condition* of privacy,[112] it has subsequently been cited to support this claim,[113] and it is therefore worth explaining why it is not convincing in this regard. The problem arises from the way in which the man on the desert island lacks privacy—granting, for the sake of argument, that he lacks privacy.

An example helps illustrate the problem. Imagine that before the man arrives on the desert island, he is on a ship with others, where he maintains the privacy of his diary. The ship then crashes and he becomes stranded alone

109. Rubel, *supra* note 93, at 278.

110. *See supra* Section IV.A.1.

111. Fried, *supra* note 29, at 482.

112. There are two reasons to think that this is not what he meant. First, while he does not clearly differentiate between normative and descriptive conceptions of privacy, his article is generally concerned with a normative conception. Second, he might not even be saying that the man on the island lacks privacy; rather, he might mean only that the man lacks what is valuable about privacy, or that privacy is meaningless on the island.

113. *E.g.*, BLANCA R. RUIZ, PRIVACY IN TELECOMMUNICATIONS: A EUROPEAN AND AN AMERICAN APPROACH 38–39 (1997). Others have also made the more general claim that a man on a desert island lacks privacy. *E.g.*, INNESS, *supra* note 10, at 44; SOLOVE, *supra* note 2, at 20; Graeme T. Laurie, *Challenging Medical-Legal Norms: The Role of Autonomy, Confidentiality, and Privacy in Protecting Individual and Familial Group Rights in Genetic Information*, 22 J. LEGAL MED. 1, 31 (2001); *see also* WESTIN, *supra* note 39, at 7 (arguing that "privacy is the voluntary and temporary withdrawal of a person from the general society").

on the island, where he is said to lack privacy. Presumably, no one would claim that at the moment he becomes stranded, he experiences a privacy loss. But if so, he does not lack privacy in the ordinary sense (i.e., in the sense that privacy has been lost); rather, it is in the sense that the concepts of privacy and limited access have become meaningless. Thus, the desert island example is not truly a case of "limited access but no privacy," and for this reason, it does not demonstrate that privacy requires both limited access and control.

To evaluate the necessity of control, we must consider cases in which control and access are theoretically possible but absent. When we do so, it becomes clear that a person can have privacy without controlling access to himself. Imagine, for example, that the government holds information about me, that no one has seen the information, and that the law prohibits the government from sharing it with anyone, even if I consent. Setting aside the question of whether I have a privacy right in this example, I think it uncontroversial to say that I have privacy in the descriptive sense. If so, personal control is not necessary for privacy.

If one does not have clear intuitions about this case, however, it might help to make a slight addition to the facts. Imagine that after a year has passed, the law changes and the government discloses some of my information. If personal informational control is necessary for privacy (such that I did not have privacy in the first year), then I would not suffer a privacy loss when my information is subsequently disclosed. But this conclusion would clearly be at odds with widespread judgements about privacy losses.

Perhaps one might try to salvage a control-based theory by giving up on the claim that the control over access needs to be in the hands of the person at issue. For example, one might argue that control can instead be in the hands of a trusted party. This modification would account for my above example involving government control (provided that the government is a trustworthy party). But it fails to save the theory, as there are clear counterexamples. Imagine, for instance, that I accidentally drop a waterproof hard drive in the ocean and that it sinks to the bottom of the ocean floor. It seems uncontroversial to say that in this case, the contents of the drive remain private, even though access to the drive is not controlled by me or anyone else.

In sum, control is neither a sufficient nor a necessary element of the condition of privacy. Like accessibility discussed above, lack of control might increase the risk of privacy losses, but it is not constitutive of them. Thus, this account of privacy does not provide a viable challenge to my core claim that privacy losses turn on access.

3. The Automation Objection

The third possible challenge to my claim that access is at the core of privacy losses comes from the idea that the automated collection and processing of personal data can cause privacy losses independently of any

human access to the data. For example, various automated technologies (such as unmonitored surveillance systems, automated web-scraping tools, and algorithms that personalize content) have been said to violate privacy rights, and one might argue that they can also cause privacy losses.

There are a few ways in which one might make this argument, but when they are analyzed, it becomes clear that the two most promising options should be rejected for reasons already identified. The first possibility is that these automated technologies can cause privacy losses by making new types of personal information accessible to human users, creating new risks of information-related harms. But this position is based on an accessibility theory of privacy loss and faces all the same problems. The second possibility is these technologies cause privacy losses by collecting and using personal data without consent. But this position relies on a control theory of privacy loss and faces all the same problems. Thus, rather than reiterate the problems with these two possible accounts of how data capture and processing could cause privacy losses, this Section will explore the possibilities that seem to remain.

One possibility is to locate the privacy losses in the act of automated data capture and its impact on people. For example, one might point to Jeremy Bentham's Panopticon to demonstrate how a surveillance system can incentivize prisoners to behave as if they are being watched, even if they are not.[114] Applied outside the prison context, what this example illustrates is how data-collecting and data-scraping technologies can—even if no person is involved—produce the same type of chilling effects on behavior and speech as actual surveillance.[115] On this basis, one might argue that mere data capture can cause a privacy loss.

However, the mere fact that automated data capture can have the same chilling effect as human access does not mean that they both cause privacy losses. It might mean that they should both be treated similarly by the law, but that is a different matter. In failing to see the difference, this account of privacy loss makes a similar error to the accessibility account, which equates access and risk of access based on the equivalence of the down-stream impacts that they might generate. An example helps illustrate the problem. Imagine, for instance, that a video camera surveillance system records the activities of someone living on a desert island. In doing so, the recording might put the person's privacy at risk, as it would take inaccessible data and make it potentially accessible (if it is possible that someone might access the recording in the future). But without relying on an accessibility-based definition of privacy loss, it is hard to see how the mere existence of the recording could be said to constitute a privacy loss.

A further problem with this account of privacy can be illustrated with a slight modification of the example. Imagine that the video surveillance

114. *See, e.g.*, Solove, *A Taxonomy of Privacy*, *supra* note 1, at 495.

115. *Cf. id.*

cameras on the desert island are turned on and capturing data, but they are not recording anything. This is a case of pure momentary data capture, without any confounding factors. To claim that a privacy loss occurs in this scenario would seem to commit one to the further view that a privacy loss occurs whenever a picture of a person is momentarily created, including by less technologically sophisticated means. But if this were correct, it would mean that even a mirror would cause a privacy loss—which presumably is not the view of those who suggest that mere data capture causes privacy losses.[116]

In order to avoid this problem, one might argue the privacy loss in fact arises from the post-capture data processing: for example, the algorithmic discovery of a person's interests, characteristics, etc. While this might seem appealing, it would entail an implausible expansion of the concept of privacy loss. It would mean, for example, that a privacy loss occurs when an electronic scale displays a person's weight or a motion detector turns on a light when a person enters a room.[117] In all of these cases, the technologies gather and process data in order to generate new information about the people interacting with them. Perhaps one might try to differentiate the electronic scale and motion detector from more personalized data processing (such as automated email analysis) on the basis of the extent to which the information is personalized. But these technologies can personalize data as well. For example, there are scales with user profiles that record and allow users to track their weight, but it is hard to imagine an argument that this technology itself causes privacy losses. They might, along with Google's email servers, increase the risk of privacy losses by making information more accessible. But as noted, I am setting aside claims here that are based in an accessibility theory of privacy, as I have rejected this approach above.[118]

This leaves one final possibility: locating the privacy loss in the interactive features of these technologies. Take, for example, the Google algorithms that "read" Gmail messages and draw inferences about the users' interests in order to provide personalized content. It could be argued that this targeting

116. Note that there is also an argument that automated systems can actually protect against privacy losses—at least relative to other permissible means of data gathering and processing. As Richard Posner has argued in the surveillance context, "computer sifting prevents most private data from being read by an intelligence officer or other human being by filtering them out." RICHARD A. POSNER, NOT A SUICIDE PACT: THE CONSTITUTION IN A TIME OF NATIONAL EMERGENCY 97 (2006); *see also* Matthew Tokson, *Automation and the Fourth Amendment*, 96 IOWA L. REV. 581, 617 (2011) ("[A]utomated systems are increasingly the means by which we *maintain* privacy in a world where virtually every transaction involves the collection of personal information.").

117. *Cf.* Tokson, *supra* note 116, at 617 ("[W]ithout some modicum of human observation, disclosure of our information to automated systems alone is ultimately no different from 'disclosure' to any other inanimate object that stores our personal data. Automated computers alone do not 'observe' us any more than a digital bathroom scale observes our weight . . . or our word-processing document observes what we type.").

118. *See supra* Section IV.A.1.

diminishes privacy.[119] This type of argument is not well-developed in the literature, but one possible claim is that the display of the personalized content causes a privacy loss. In support of this claim, one might draw a comparison to intrusion upon seclusion, as the receipt of personalized content might be as unsettling as an intrusion (especially if it reveals interests that one thought were secret, or interests that one had not recognized in oneself). But this would vastly expand the concept of privacy losses, capturing cases of non-human intrusions clearly unrelated to privacy. As Matthew Tokson highlights: "It is easy to anthropomorphize these capable machines, to think of the computers that analyze personal data and send targeted advertisements as the equivalent of a human salesman tailoring his sales pitch to his audience."[120] But clearly they are not.

In sum, while there is intuitive appeal to the claim that data capture and processing technologies can cause privacy losses themselves (i.e., independently of access to the data by someone), this position does not withstand scrutiny. The strongest arguments in support of it rely on accessibility-based or control-based theories of privacy loss, which should be rejected for the reasons identified above.[121] Further, these are not the only problems with this position, for as is discussed next, there are epistemic criteria for privacy losses that these technologies will also fail to satisfy.

B. EPISTEMIC MERIT

My argument thus far—that access is a requirement of privacy losses—raises the question of whether any type of access is sufficient, or whether it must meet additional epistemic criteria, such as those of knowledge. Although knowledge is included in many limited-access theories of privacy (defining privacy in terms of limits on knowledge) and control-based theories (defining privacy as control over knowledge), this element of their definitions has received little attention and there is almost no literature on it.[122]

119. *Cf.* Fallis, *supra* note 97, at 165 (arguing that privacy is diminished by automated targeted advertising).

120. Tokson, *supra* note 116, at 616–17.

121. *See supra* Sections IV.A.1, IV.A.2. It is also worth flagging here that on the distinct matter of privacy violations—to which I will turn in Part V—courts have held that human access is necessary for a Fourth Amendment violation. *See, e.g.*, United States v. Karo, 468 U.S. 705, 712 (1984) ("The mere transfer to Karo of a can containing an *unmonitored* beeper infringed no privacy interest. . . . To be sure, it created a *potential* for an invasion of privacy, but we have never held that potential, as opposed to actual, invasions of privacy constitute searches for purposes of the Fourth Amendment." (emphasis added)); *see also* POSNER, *supra* note 116, at 97 (arguing that automated sifting of data "is neither a search within the meaning of the Fourth Amendment nor 'surveillance' within the meaning of FISA"); Orin S. Kerr, *Searches and Seizures in a Digital World*, 119 HARV. L. REV. 531, 551 (2005). *See generally* Tokson, *supra* note 116, at 609–19 (arguing that when personal information is exposed only to automated systems, no loss of privacy occurs).

122. *See supra* notes 96–97 and accompanying text.

The intuition that knowledge might be required to cause a privacy loss is not difficult to generate. Imagine, for example, that I have a dream in which I learn a secret piece of information about a real person; that I later forget that I learned this in a dream and so believe that it is true; and that the information happens to be true as a matter of mere chance. It seems uncontroversial to suggest that the person I dream about does not suffer a loss of privacy via my dream—and that the reason for this is that my belief about him is not connected in any way to the fact that makes it true. It is true merely as a matter of luck.

The idea that mere true beliefs (i.e., beliefs that are true by luck) do not count as knowledge dates back at least as far as Plato, but the task of identifying the additional epistemic criteria that must be satisfied has challenged philosophers for generations.[123] Increasingly intricate theories have been proposed in a vast literature, and there is still no consensus. Luckily, it is possible to answer the question of whether knowledge is required for privacy losses without reaching a definitive answer on the question of what counts as knowledge. An understanding of some of the foundational approaches is sufficient to clarify whether knowledge—or as I will propose, a set of epistemic desiderata related to knowledge—is an essential element of a privacy loss.

1. Theories of Knowledge

A canonical way of explaining why merely true beliefs do not count as knowledge is to impose a *justification* requirement for knowledge.[124] This is known as the "justified true belief" ("JTB") theory of knowledge and requires that a true belief be adequately grounded in evidence and reasons to count as knowledge.[125] This definition has clear intuitive force and was for a long time widely accepted. But it is now generally rejected for failing to exclude cases of epistemic luck.

123. For a short overview of this history and the various criteria that have been proposed, see generally Jonathan Jenkins Ichikawa & Matthias Steup, *The Analysis of Knowledge*, *in* THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., Summer 2018 ed.), *available at* https://plato.stanford.edu/archives/sum2018/entries/knowledge-analysis [https://perma.cc/JD9V-28WT].

124. The term "justification" is sometimes used in two different ways, which can cause some confusion. The term is traditionally used to refer to when a person has good reasons/evidence for a belief and is able to identify those reasons and evidence. However, the term is also occasionally used to refer to other ways in which a belief might be epistemically warranted (e.g., the types of warrant identified by the "causal" and "reliabilist" theories discussed below). In the interest of clarity, I will use the term in its narrower traditional sense.

125. There are different views on the question of how this line of reason-giving must ultimately be grounded for a true belief to count as knowledge, but such details are not crucial here. One view is that the process must end in some "foundational reasons" that are not supported by other reasons (a position known as "foundationalism"); another view is that the process can be grounded in a system of mutually supporting beliefs that cohere (a position known as "coherentism").

The problem with the JTB theory can be illustrated with the following example—a modification of the facts of the Supreme Court case *Kyllo v. United States*, which I will explore in depth in Part V:

> *False Premise*: A police officer uses an infrared thermometer to measure the heat being emitted from the walls of D's house (where high-voltage lamps are being used to grow marijuana). The thermometer indicates that it is 85 degrees inside. The officer has good reason to believe that the thermometer is accurate, but it is actually broken and always gives a reading of 85. However, as a matter of pure coincidence, it is 85 inside the house, so the police officer forms a true belief about the temperature.[126]

Setting aside the question of whether the police officer has caused D a privacy loss (or a privacy violation)—to which I will later return—my point here is about the epistemic status of his belief. Even though he has a justified true belief about the temperature in the house, his belief is true by pure luck, so he cannot be said to have knowledge of the temperature.

This type of case of epistemic luck (known as a "Gettier problem") spurred an attempt to find additional criteria to add onto the JTB theory of knowledge. For example, some thought that it might be saved by adding a "defeasibility criterion"—i.e., a requirement that the justification not be undermined by the addition of new facts (such as the fact that the infrared thermometer was broken in my example).[127] But this and other additional criteria proved unable to remedy all of the problems with it.[128]

This led many philosophers to reject the traditional approach of defining knowledge internally (in terms of reasons/evidence), and to instead ground knowledge externally. An early example of this approach is the "causal theory" of knowledge, which suggests that a true belief is knowledge only when it is caused by the fact that makes it true. By imposing the causation requirement,

126. This type of counter-example to the JTB theory of knowledge—in which a justified true belief is inferred from a justified false belief, so is true merely by luck—is often referred to as a "Gettier problem," based on the work of Edmund Gettier. *See generally* Edmund Gettier, *Is Justified True Belief Knowledge?*, 23 ANALYSIS 121 (1963).

127. The basic idea behind the defeasibility criterion is that knowledge does not turn solely on the evidence that one possesses (and the reasons for the belief that they provide), but also on the evidence that one does not possess. In order to have knowledge, there cannot be any evidence that would, if one possessed it, undermine the justification of one's belief. This criterion is often attributed to Keith Lehrer & Thomas Paxson, Jr. *See* Keith Lehrer & Thomas Paxson, Jr., *Knowledge: Undefeated Justified True Belief*, 66 J. PHIL. 225, 229–31 (1969).

128. *See generally* Marshall Swain, *Defeasibility Theory of Knowledge*, *in* ROUTLEDGE ENCYCLOPEDIA OF PHILOSOPHY 276 (1st ed. 2016) (discussing the problem of misleading defeaters). A related earlier proposal was to add a "no false lemmas" criterion, which could deal with the problem posed by *False Premise*, but it soon became apparent that it was insufficient for reasons identified by Alvin I. Goldman. *See* Alvin I. Goldman, *Discrimination and Perceptual Knowledge*, 73 J. PHIL. 771, 773–75, 786–88 (1976) (proposing the "barn facsimiles" hypothetical).

the theory excludes many beliefs that are true merely by luck. For example, it correctly excludes Gettier-style cases such as *False Premise*, where the actual temperature in the house has no causal connection to the officer's true belief. But it soon became clear that this theory was also too permissive in some cases, such as in this modification of the facts:

> *Unreliable Process*: A police officer's infrared thermometer indicates that it is 85 degrees inside D's house. The thermometer is unreliable and only works one percent of the time, but the police officer does not know this and so he believes it. In this case, he is lucky and it is working, so he ends up forming a true belief about the temperature inside the house.

This belief satisfies the requirements of the causal theory of knowledge, as there is a causal connection between the temperature in the house, the measurement by the thermometer, and the officer's true belief. But as above, his belief is true by mere luck and so cannot be considered knowledge. In this type of case, the causal criterion is insufficiently restrictive (and in a way that is avoided by the defeasibility criterion, rejected above). In other cases, however, it turns out that it is overly restrictive. For example, the causal theory does not allow for knowledge of *a priori* propositions (e.g., "85 is a larger number than 80"), counterfactuals (e.g., "if he had not been using heat lamps in his house, it would not have been as hot inside"), or other beliefs that are logically or mathematically true, as these beliefs are not causally connected to the facts that make them true.

In order to address the limits of the causal theory, philosophers developed the third and final theory of knowledge that I will discuss: reliabilism. Instead of requiring an appropriate causal connection between the fact and the belief, reliabilism requires that the process resulting in the belief produces true beliefs sufficiently often.[129] Like the causal theory, the reliabilist theory provides a fairly straightforward account of what makes knowledge non-accidental. But unlike the causal theory, it can explain why the police officer lacks knowledge when his true belief is based on an unreliable thermometer that rarely works.

Like the other theories, however, reliabilism also has limits that render it an insufficient theory of knowledge. One problem is illustrated by this variation on a core example from the literature:

> *Mechanistic Belief*: A police officer correctly believes that it is 85 degrees inside D's house because he has a chip in his brain that is connected to an accurate thermometer inside the house. He does not know, however, that he is connected up to the thermometer in

129. For an excellent overview of reliabilist theories and their critics, see Alvin Goldman & Bob Beddor, *Reliabilist Epistemology*, *in* THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., Winter 2016 ed.), *available at* https://plato.stanford.edu/archives/win2016/entries/reliabilism [https://perma.cc/636X-9HZC].

> this way. He just experiences a strong belief about what the temperature is.[130]

In this case, his true belief satisfies the basic requirement of the reliability theory, as his belief-formation process is highly reliable. But it would be odd to say that he has knowledge of the temperature, as he does not have any reason to trust his beliefs. In this way, the reliabilist theory is (like the others) insufficiently restrictive, which is just one of several grounds on which it has been rejected.[131]

2. Epistemic Warrant and Privacy

In light of the challenges facing these and other theories of the nature of knowledge (I have only summarized the foundational attempts), some philosophers have argued that we should adopt a pluralistic approach to knowledge. For example, William Alston argues that we should abandon the attempt to find a single property of beliefs that is picked out by the term "justified" and instead recognize that there are many different "ways in which beliefs can be better or worse from an epistemic point of view."[132] These "epistemic desiderata" include many of the criteria that have been identified in the search for a unified theory of knowledge (e.g., reliability, adequate grounding in evidence, etc.), but on the pluralistic approach, none is taken to be foundational to knowledge.[133]

Setting aside the question of whether pluralism is ultimately satisfying as an approach to defining the nature of knowledge (a question that is outside the scope of this paper), I will suggest that this approach can be adopted to define the epistemic criteria that must be satisfied for privacy losses. To explain how, it is necessary to first take a step back and make two preliminary observations regarding the scope of the epistemic criteria that must be satisfied for a privacy loss to occur.

First, at one end of the spectrum, it is clear that a completely unwarranted belief about a person is not sufficient to diminish that person's privacy, for

130. *Cf.* KEITH LEHRER, THEORY OF KNOWLEDGE 163–64 (1990) (developing the widely-discussed example of "Mr. Truetemp").

131. There are four other dominant critiques. The first problem is that reliability does not appear to be *necessary* for a true belief to be warranted. *See* Stewart Cohen, *Justification and Truth*, 46 PHIL. STUD. 279, 283 (1984). The second is the problem of defining the level of generality at which the relevant process is defined. *See* Richard Feldman & Earl Conee, *Internalism Defended*, 38 AM. PHIL. Q. 1, 3 (2001). The third is the problem of easy knowledge. *See* Jonathan Vogel, *Reliabilism Leveled*, 97 J. PHIL. 602, 603 (2000). The fourth is the problem of explaining why, on a reliabilist theory, knowledge is more valuable than mere true belief. *See* Linda Zagzebski, *The Search for the Source of Epistemic Good*, 34 METAPHILOSOPHY 12, 12–20 (2003).

132. *E.g.*, WILLIAM P. ALSTON, BEYOND "JUSTIFICATION": DIMENSIONS OF EPISTEMIC EVALUATION 21 (2005); William P. Alston, *Epistemic Desiderata*, 53 PHIL. & PHENOMENOLOGICAL RES. 527, 527 (1993).

133. For more on the "epistemic desiderata" that Alston proposes, see generally ALSTON, *supra* note 132.

reasons I discussed at the start of this Section. The hypothetical of the dream illustrated that in order to cause a privacy loss, personal information must be accessed in a way that has some epistemic merit. The belief cannot be true by mere luck.

Second, at the other end of the spectrum, a review of the various theories of knowledge suggests that privacy losses can occur without knowledge —however it is defined. For example, justified true beliefs are not needed, as Anita Allen has argued. She illustrates this point with the example of a person who reads a celebrity's diary that is known to contain a mix of "saucy facts and fantasy."[134] In this case, it seems clear that the celebrity can experience a loss of privacy with respect to some of the information in the diary, even though the reader of the diary cannot be justified in believing any particular statement in the diary (given that the reader knows that any statement could be fantasy). This suggests that true beliefs need not be justified to cause privacy losses—nor do they need to satisfy more restrictive versions of this approach, such as those that add defeasibility criteria.[135]

Likewise, it is clear that a causal connection is not needed to cause a privacy loss. Imagine, for example, that I have strong evidence that all the members of a given group (e.g., the group of people who voted for a particular political candidate) share a particular trait (e.g., each member has an income of over a million dollars). If I meet someone who tells me he is a member of the group, I will be able to infer that he has that trait, and if this is true, it seems clear that he will have lost privacy in this fact. However, the fact that he has the trait will not have caused my true belief. Rather, I will have inferred this fact through the application of basic deductive logic.[136] Thus, the causal theory does not need to be satisfied for a privacy loss to occur.

The same is true of reliabilist theories of knowledge, as a privacy loss can occur even if information is accessed in a way that is not reliable. Don Fallis

134. ALLEN, *supra* note 10, at 21. Allen acknowledges that the "justified true belief" is just one way of defining knowledge.

135. For example, imagine that I correctly believe that my friend attended a secret club because I saw him entering it, but that unbeknownst to me, he has an identical twin brother who also attended the club. As knowledge of this additional fact would defeat the justification for my belief, I would not "know" that my friend attended the club. But clearly, he would have lost privacy in this fact.

136. My deductive inference takes the standard form: all Xs are Y; P is an X; therefore P is Y. The most famous example of this is: all men are mortal; Socrates is a man; therefore Socrates is mortal. This fails to satisfy the causal theory in two ways. First, it relies on a generalized belief ("all Xs are Ys") that itself fails to satisfy the causal theory. *See generally* JONATHAN DANCY, INTRODUCTION TO CONTEMPORARY EPISTEMOLOGY 34 (1985) (explaining that the belief "all men are mortal" is not caused by the fact that all men are mortal). Second, the conclusion relies on logical relationships, which are generally not thought to be causal relationships. While Alvin Goldman proposed a sophisticated version of the causal theory that was meant to work around these problems, this faces problems of its own. *See generally* Peter D. Klein, *Knowledge, Causality, and Defeasibility*, 73 J. PHIL. 792 (1976) (arguing that a causal theory cannot provide an adequate explanation of inferences).

illustrates this with the example of a compulsive liar who knows a secret about someone and discloses it to others who know he is a compulsive liar and thus have good reason to doubt him (even though in this case he happens to be telling the truth).[137] Fallis argues that in this case, the disclosure causes privacy loss even though the source is unreliable (and thus cannot support knowledge under a reliabilist account) and is known to be unreliable (and thus cannot support knowledge under a justification-based account).[138]

If the same is true for every other theory of knowledge when considered in isolation—which seems plausible, though not necessary to verify for reasons that will soon become clear—one might conclude that something more than true belief, but less than knowledge, is necessary to cause privacy losses. This is the conclusion reached by Don Fallis, who is one of the few philosophers who has devoted significant attention to this question.[139] On the basis of this conclusion, he attempts to identify the type of connection between belief and fact that is sufficient to cause a privacy loss. Drawing on causal theories of knowledge, he argues that certain types of causal connections will be sufficient, though not necessary.[140]

While this analysis by Fallis offers valuable insights, it is limited by the fact that he treats each theory of knowledge individually. To identify the necessary conditions for privacy losses, it is helpful to look past theories of knowledge in isolation.

If they are instead seen collectively—i.e., as each identifying one criterion in a broader set of epistemic desiderata—a more significant insight emerges: that in order to cause a privacy loss, the access must satisfy at least one desideratum in this set. This can be seen in all of the examples discussed above (each of which seemed to indicate that knowledge was not necessary). Take, for example, Anita Allen's example of the celebrity diary. In this case, the justification criterion for knowledge is not satisfied, but the causal criterion is, as there is a causal connection between the truth of the saucy facts, the author's inclusion of them in his diary, and the reader's belief in them.[141] In Fallis' example, by contrast, the justification and reliabilist criteria are not satisfied, but the causal criterion is, as there is a causal connection between the truth of the information, the liar's knowledge of it, and the listener's belief in it. Conversely, in my example of inferred traits, the causal criterion is not satisfied, but the justification criterion is (given that the belief is based on valid reasons), as is the reliabilist criterion (given that a deductive inference from true premises is a reliable means of producing true beliefs).

137. Fallis, *supra* note 97, at 157–60.

138. *Id.*

139. *See id.* at 160.

140. *Id.* at 160–61.

141. In addition, the reliabilist criterion is satisfied, assuming that reading people's diaries is generally a reliable way of learning information about them.

Furthermore, this pluralist account of the necessary conditions can explain cases in which privacy is not lost, such as in the case of true beliefs formed via dreams. These true beliefs are not justified, caused by the facts that make them true, generated through a reliable process, etc., and so are properly excluded from causing privacy losses.

For ease of reference, I will refer to this requirement of satisfying at least one epistemic desideratum in the set—but failing to satisfy all the criteria necessary for knowledge—as the requirement of "epistemic merit."

C. TRUTH

Thus far in my argument that privacy losses turn on access, I have implicitly assumed that true information must be accessed. But this assumption must be interrogated, as there are several privacy scholars who have argued that privacy losses can occur through the acquisition and disclosure of falsehoods.[142]

In support of the argument that privacy losses can occur through the acquisition of false information, Pierre Le Morvan cites Anita Allen's example (discussed above) of a celebrity diary containing "saucy facts and fantasy" being covertly obtained and read by a stranger.[143] While Allen offers the example to support the claim that privacy losses do not depend on knowledge, Le Morvan uses it to argue that privacy losses do not even depend on true beliefs.[144] He argues that if the stranger "believes several fantastical entries to be true," the celebrity has "incurred a loss of privacy relative to these fantastical entries even though they are false."[145] There are confounding factors in this example, however, that limit its value in exploring whether access to false information can cause a privacy loss. Most significantly, the stranger may have violated the celebrity's privacy rights in gaining access to the diary, which may confuse our intuitions about whether there was also a privacy loss. In addition, the stranger in the case does gain some true knowledge about the contents of the celebrity's diary, even if it leads to some false beliefs about the celebrity's life.

Without these elements, it is hard to imagine that anyone would think the celebrity's privacy is implicated in this case. For example, imagine that the facts of the scenario generally stay the same, except for the following two changes. First, the stranger finds the diary in a park, so there is no privacy violation. Second, the stranger is mistaken in thinking he has the celebrity's diary; in fact, he has the diary of the celebrity's friend, on the basis of which he develops the same false beliefs about the celebrity. If access to false

142. *E.g.*, Johnson, *supra* note 2, at 162; Pierre Le Morvan, *Privacy, Secrecy, Fact, and Falsehood*, 40 J. PHIL. RES. 313, 316–21 (2015); Rubel, *supra* note 93, at 277.

143. ALLEN, *supra* note 10, at 21; Le Morvan, *supra* note 142, at 318.

144. Le Morvan, *supra* note 142, at 318.

145. *Id.*

information about a person could cause a privacy loss, we would be forced to conclude that the privacy of the celebrity (and not her friend) would be diminished here. This is implausible. If anyone could be said to experience a privacy loss in this case, it would surely be the person whose diary has been read.

One might argue, however, that our views about this case are influenced by the fact that the thief could have easily discovered he had the wrong diary and that the information was false. Perhaps this epistemic consideration is relevant. For example, one might seek to defend the claim that access to false information can cause a privacy loss by limiting it to cases in which there is no way to know the information is false (i.e., where the belief in it is unfalsifiable). Imagine, for example, that a doctor performs a test that incorrectly indicates that a patient has a rare genetic disease. There is no reason to doubt the accuracy of the test, and no way to discover that the patient does not have the disease. One might argue that in this case, the patient experiences a loss of privacy by virtue of the doctor's access to this false information about him.

Even if limited in this way, however, the claim that access to false information can cause a privacy loss has untenable implications regarding retroactivity and the correction of false beliefs. While advocates of this account of privacy loss have not addressed the question of what happens when false beliefs are corrected, it seems that they would need to adopt one of two possible positions, neither of which is plausible. One possibility is that access to false information creates a privacy loss, and the subsequent realization that the information is false does not modify this loss. It is hard to imagine anyone advancing this view. The other possibility is that access to false information creates a privacy loss, and that this privacy loss ceases to exist when the mistake is realized. This seems more plausible, but it would mean that privacy losses turn entirely on what people think they know, which is incompatible with the project of identifying the criteria that define the condition of privacy as a state of affairs.[146] The loss of the condition of privacy is like the loss of any other state of affairs; whether there has been a loss is a fact about the world, which does not depend on what anyone believes about it. In the case of the doctor and the genetic test, the best way to describe what has happened is that the patient thought that he had experienced a privacy loss, but in fact, he had not.

Thus far, my analysis of privacy losses and truth has focused on cases of data acquisition, but it is worth exploring whether there is anything different about data disclosure, as challenges to my position are often supported with

146. However, it might be compatible with an alternative, phenomenological, conception of privacy. *Cf.* Parker, *supra* note 42, at 278 (differentiating between five separate questions: "(1) whether a person *has* lost or gained privacy, (2) whether he *should* lose or gain privacy, (3) whether he *knows* that he has lost or gained privacy, (4) whether he *approves or disapproves* of the loss or gain, and (5) how he *experiences* that loss or gain").

examples involving the disclosure of false information. For example, Alan Rubel offers this hypothetical:

> Suppose that a healthcare provider confuses medical records such that P's name is attached to the medical history of another. If the provider releases that record, it would seem that P's privacy has diminished. P has a legitimate complaint against the medical provider and that complaint is grounded in a diminution of her privacy.[147]

While I imagine most would agree that P has a legitimate complaint in this hypothetical, Rubel's claim that this complaint is grounded in a loss of privacy is more controversial and must be analyzed. The potential problem is that P's complaint may instead be grounded in a related type of harm. According to Parent, for example: "The spreading of falsehoods or purely subjective opinions about a person does not constitute an invasion of his privacy. It is condemnable in the language of libel or slander."[148] Building on this claim, one might argue that the intuition that privacy is implicated in Rubel's scenario is motivated by the fact that the publication of true and false information can be equally harmful, but that it is a mistake to think all information-related harms arise from privacy losses.

This line of objection strikes me as compelling, but it is unclear how far it applies. The open question is whether information-related harms (such as those captured by libel and slander) are at the core of all legitimate complaints about access to false information. Rubel argues that they are not. In support of this view, he asks that we imagine that in his hypothetical, the release of the false medical record *benefits* P. According to Rubel, P would still have a valid complaint: "P's complaint is that the record was released, period—and that release diminishes her privacy."[149] Again, I imagine many will share the view that P has a legitimate complaint in this case, even if the information is beneficial. But it would be a mistake to conclude, on the basis of this view, that access to false information can itself cause a privacy loss. There are two reasons for this.

The first reason is that Rubel (like others who argue that informational conceptions of privacy are insufficient)[150] is making a claim about whether P has a privacy-based complaint. Even if this claim is correct, however, it does not tell us whether P has suffered a privacy loss. Rather, it tells us whether she has suffered a privacy violation, and P can suffer a privacy violation without suffering a privacy loss.

147. Rubel, *supra* note 93, at 277*; see also* Johnson, *supra* note 2, at 162 (discussing falsehoods and privacy).

148. Parent, *supra* note 6, at 269 n.1.

149. Rubel, *supra* note 93, at 277.

150. *E.g.*, Gavison, *supra* note 10, at 433; Johnson, *supra* note 2, at 160–61.

The second problem with Rubel's conclusion is that an aspect of P is in fact accessed when her name is published along with false information about her. Her situation is similar to that of a naked person who is seen through her fence by a Peeping Tom who develops the false belief that she has a tattoo (and gains no other true information about her). This person clearly suffers a privacy loss, but not by virtue of the false belief; rather, it is because the Peeping Tom was in fact looking at her. Likewise, when P's name is published along with false information about her, she experiences a privacy loss (and also, possibly, a privacy violation) by virtue of the public attention to her that the publication creates.

Finally, if it were true that privacy losses could occur through the disclosure of false information, people would not be able to maintain their privacy by disclosing false information about themselves to others. Imagine, for example, that I tell a lie about my location in order to protect my privacy and prevent someone from knowing where I am. If false information could cause privacy losses, I would experience a privacy loss by virtue of my lie. This is an implausible description of what happens when people tell lies about themselves.

Perhaps an advocate of this account of privacy loss would respond that such lies should instead be seen as creating a trade-off between privacy losses—for example, that when I lie about my location, I lose privacy with respect to my false location and maintain privacy with respect to my true location. But this would still involve the implausible suggestion that I can lose privacy in a lie about myself. The more natural description is that I do not experience any privacy loss in this case—that my lie maintains the privacy of my location completely. This is yet another reason to reject the claim that access to false information can itself cause a privacy loss.

V. PRIVACY VIOLATIONS

I have thus far argued that on the best account of the descriptive concept of privacy used in ordinary language, a loss of privacy occurs when true information about a person is accessed by another in a way that has epistemic merit. In what follows, I will build on this account, arguing that privacy rights should be understood as restricting the means by which privacy losses can legitimately occur—and thus that a person suffers a privacy violation when a restriction on the permissible means of obtaining this type of access is breached.

To be clear, this is an argument about the general nature of privacy violations, not their precise content, so I do not take a position on the question of which means of access are impermissible or what types of facts should be protected. Because my account unifies privacy along its descriptive rather than its normative dimension, it is compatible with significant disagreement on these and other related normative questions.

At the same time, however, the theory of privacy that I develop has a critical edge,[151] challenging widespread claims about whether and how privacy rights are violated by the aggregation, unconsented use, and inference of personal information. Paying attention to the loss/violation distinction reveals that the scholarship on these issues has misinterpreted key Supreme Court cases, including the landmark technology cases *Carpenter v. United States*[152] and *Kyllo v. United States*.[153] It also helps clarify potential problems with expanding privacy rights in ways that have been suggested.

Note that in exploring the lessons to be learned from my account of privacy violations, I will primarily focus on those that follow from its grounding in an access-based account of privacy losses, as these are likely to be the most controversial. But the other two criteria identified above —epistemic merit and truth—also highlight important questions that require further attention. I will briefly explore these in my analysis of the legal status of inferences, identifying the unrecognized role of these two criteria in Fourth Amendment violations.

Finally, it is worth preemptively clarifying one important aspect of my account of the relationship between privacy losses and violations. While I define privacy violations in terms of privacy losses (which provides for the coherence of privacy rights), this does not mean that a privacy violation can only occur if privacy is in fact lost. Rather, a restriction on a means of access that is meant to protect against privacy losses can be violated even if access is not ultimately achieved.[154]

A. THE PATH-BASED ELEMENT

The centrality of the means of access in privacy violations has received little explicit attention in the literature, but upon analysis, it is clear that

151. *Cf.* DWORKIN, *supra* note 15, at 116–18 (explaining that the theory that best fits and justifies an area of law will often not fit all of our judgements about it; rather, to achieve coherence, it will often find some judgements to be mistaken or misconceived).

152. Carpenter v. United States, 138 S. Ct. 2206, 2211 (2018).

153. Kyllo v. United States, 533 U.S. 27, 29 (2001).

154. This is in line with the decisions in *Amati v. City of Woodstock*, 829 F. Supp. 998, 1010 –11 (N.D. Ill. 1993), *aff'd*, 176 F.3d 952 (7th Cir. 1999) ("The court is of the opinion that if Illinois courts were to adopt this tort, they would also recognize a cause of action as pled in this case. Accordingly, the lack of allegations that anyone actually listened to the recorded telephone conversations do not defeat plaintiffs' claims."); *Harkey v. Abate*, 346 N.W.2d 74, 76 (Mich. Ct. App. 1983) ("In our opinion, the installation of the hidden viewing devices alone constitutes an interference with that privacy which a reasonable person would find highly offensive. And though the absence of proof that the devices were utilized is relevant to the question of damages, it is not fatal to plaintiff's case."); and *Hamberger v. Eastman*, 206 A.2d 239, 242 (N.H. 1964) (holding that, in spite of the fact that the tenants did not allege the landlord actually utilized the listening device, their complaint adequately stated an action for invasion of privacy). *See also* Schoeman, *supra* note 6, at 4 ("We can also envision situations in which we would want to say that a person has not in fact suffered loss of privacy but has suffered a violation of his right to privacy.").

privacy rights do not restrict access per se. Rather, they restrict specific means of access, which I will refer to as "path-based" restrictions for short.

This feature of privacy rights can be found across the constitutional, statutory, and common law rights that restrict the acquisition of personal information. For example, the Fourth Amendment right protecting reasonable expectations of privacy does not protect reasonable expectations that a given piece of information will not be accessed (i.e., it is not a right against mere privacy losses); rather, it protects reasonable expectations that the information will not be accessed in certain ways. The literature has often overlooked this feature of the Supreme Court's *Katz* jurisprudence (giving rise to some confusion discussed in the next Sections),[155] but it is hard to imagine anyone rejecting my claim. For example, as *United States v. Jones* highlights, it is impermissible to track a person's movements with GPS for an extended period, but permissible to do so by following them in a car, even if both methods reveal the same location data.[156] Even information in the home is protected by path-based restrictions, as the Court explains in *Kyllo*.[157] This feature of privacy rights can also be seen in state and federal privacy statutes, which do not restrict access per se, but rather specific means of access, such as wiretapping telephones, intercepting electronic communications, using two-way mirrors in specified areas, and looking through home windows.[158] Likewise, the common law tort of intrusion on seclusion only provides a cause of action if the means of access is highly offensive to a reasonable person.[159]

The path-based character of privacy rights can also be seen in the various sources of privacy law that restrict the disclosure of information. For example, the constitutional right to information privacy, which has been assumed (though not formally recognized) by the Supreme Court,[160] is path-based. In *Whalen v. Roe*,[161] the Court held that the Constitution might protect an individual interest in avoiding disclosure of personal matters, but that this

155. *See* sources discussed *infra* Sections V.B.1, V.C.1.

156. United States v. Jones, 565 U.S. 400, 404 (2012).

157. The Court frames the point from the other direction, but its substance is the same: "The police might, for example, learn how many people are in a particular house by setting up year-round surveillance; but that does not make breaking and entering to find out the same information lawful." *Kyllo*, 533 U.S. at 35 n.2.

158. *See* Solove, *A Taxonomy of Privacy*, *supra* note 1, at 491–93.

159. RESTATEMENT (SECOND) OF TORTS § 652B (AM. LAW INST. 1977). For example, this has been interpreted as providing a right against certain forms of surveillance in public in *Nader v. General Motors Corp.*, 255 N.E.2d 765, 770–71 (N.Y. 1970), photography in a medical office in *Estate of Berthiaume v. Pratt*, 365 A.2d 792, 793 (Me. 1976), and audio recording in a bedroom in *Hamberger*, 206 A.2d at 242.

160. In two cases, the Supreme Court has stated that it would assume there is such a right for the sake of its analysis, but found that it would not have been violated. NASA v. Nelson, 562 U.S. 134, 138–39 (2011); Whalen v. Roe, 429 U.S. 589, 599–600 (1977). In *NASA v. Nelson*, Justice Scalia strongly objected to the Court's working assumption, declaring: "A federal constitutional right to 'informational privacy' does not exist." *Nelson*, 562 U.S. at 160 (Scalia, J., concurring).

161. *Whalen*, 429 U.S. at 603–04.

would only require the state to take measures to avoid unreasonable disclosure. As the state had taken adequate security measures (i.e., it had restricted the pathways by which disclosure might occur), there was no violation of the right.[162] Path-based restrictions—regarding how and to whom information is disclosed—can also be found in a wide range of federal privacy laws.[163] These laws also generally impose related duties, such as data security requirements, which likewise restrict means of access.

In sum, across the various areas of privacy law that restrict the acquisition and disclosure of personal information, privacy violations are path-based. The question is not *whether* information has been accessed, but rather *how* it has become accessible. (Under many of these areas of law, a violation also depends on *what* information is at issue, but that is not relevant here).

Of course, as a normative matter, one might argue there is no reason why privacy rights should be limited to path-based restrictions. For example, George Brenkert has argued that the acquisition of certain *types* of information can constitute a privacy violation regardless of the way in which the acquisition occurs:

> [T]here are certain things which people (in their various roles as employers, government officials, physicians, etc.) and institutions (governments and businesses, etc.) ought not to know about individuals, however they might come to know these facts . . . . For example, it would be wrong, however they went about it, for government officials to make it their business to know the details of the sexual practices of each particular citizen.[164]

Brenkert further argues that "since they ought not to know such facts, those individuals who are the ultimate object of this knowledge may legitimately object to a violation of their rights."[165]

While many might agree with Brenkert that it would be wrong for government officials to seek to discover their citizens' sexual practices regardless of the means, it does not necessarily follow that privacy losses can (in and of themselves) constitute violations of privacy rights. Two aspects of this conclusion require unpacking. First, even if one agrees with Brenkert that

162. *Id.* at 605–06. Some lower courts have gone further and required the government adopt the least intrusive means of disclosure. *See, e.g.*, Donohue v. Hoey, 109 F. App'x 340, 361 (10th Cir. 2004).

163. *See, e.g.*, Privacy Act of 1974, 5 U.S.C. § 552(a) (2012); Right to Financial Privacy Act of 1978, 12 U.S.C. §§ 3401–3422 (2012); Gramm–Leach–Bliley Act, Pub. L. No. 106-102, § 1(a), 113 Stat. 1338 (codified as amended in scattered sections of 12 U.S.C. and 15 U.S.C.); Video Privacy Protection Act of 1988, 18 U.S.C. § 2710 (2012); Driver's Privacy Protection Act of 1994, 18 U.S.C. §§ 2721–2725 (2012); Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, § 1(a), 110 Stat. 1936 (codified as amended in scattered sections of 29 U.S.C. and 42 U.S.C.).

164. George G. Brenkert, *Privacy, Polygraphs and Work*, 1 BUS. & PROF. ETHICS J. 19, 20 (1981).

165. *Id.*

the government official who sees the person engaged in sexual activity "ought not to know such facts" about the person, it does not follow that he has a duty to not know them (and violates the person's rights merely by knowing them). Second, the government officials in Brenkert's hypothetical have intent, which seems to drive the intuition that they are behaving wrongfully. To see this, imagine that there is no intent: for example, that a government official walks into a public bathroom where a person is engaged in sexual activity. In this case, it is hard to imagine a plausible normative argument that such an official violated the person's rights.

Thus, it seems that most would agree with Judith Thomson when she writes:

> [N]one of us has a right over any fact to the effect that that fact shall not be known by others. You may violate a man's right to privacy by looking at him or listening to him; there is no such thing as violating a man's right to privacy by simply knowing something about him.[166]

Privacy rights that impose access restrictions should—as a positive and normative matter—be interpreted as restricting the means of access, not access itself.

B. DATA AGGREGATION AND USE

While my argument thus far has focused on privacy rights that restrict access, there are practices that are widely said to violate privacy rights without violating access restrictions: namely, the aggregation and unconsented use of personal data.[167] If it were true that these practices violated privacy rights in and of themselves (i.e., independently of violating any access restrictions), this would pose a problem for my account of privacy violations, along with the account of privacy losses that underlies it. As I will argue, however, they do not. Differentiating between privacy losses and violations reveals the nature of the mistake and sheds light on how we should actually understand the privacy implications of the aggregation and unconsented use of personal data.

1. No Right Against Aggregation

Both advocates and critics of a privacy right against data aggregation often agree on one point: that courts have rarely recognized such a right, but that the Supreme Court did so in the recent landmark case *Carpenter v. United States*,[168] as well as in *United States Department of Justice v. Reporters Committee for*

166. Thomson, *supra* note 33, at 307; *see also* Scanlon, *supra* note 33, at 315 (describing privacy rights as enforcing "norms specifying when, where, and in what ways we may and may not be observed, listened to, questioned, and in other ways kept track of").

167. For an overview of these positions in the literature, see Solove, *A Taxonomy of Privacy*, *supra* note 1, at 505–11.

168. Carpenter v. United States, 138 S. Ct. 2206, 2208 (2018).

Freedom of the Press.[169] This point of agreement is, however, based on a mistaken reading of the two cases. Attention to the loss/violation distinction reveals the mistake and clarifies what the Court actually held.

In *Carpenter*, the Court addressed the question of whether the government had violated the defendant's reasonable expectations of privacy under the Fourth Amendment when it obtained historical "cell-site location information" ("CSLI") data from his wireless carriers.[170] This included 12,898 location points over a period of 127 days.[171] Under the Court's well-established "public exposure" doctrine, it seemed that this data would be excluded from Fourth Amendment protection. But in a significant shift, the Court found this doctrine inapplicable, in part because of the aggregated nature of the data.[172] The Court explained that although Carpenter had exposed each of his physical movements to different people at different places and times, he had not exposed the whole of his physical movements to any single person.[173] In other words, by reframing the doctrinal question to focus on aggregated data and a single observer (rather than disaggregated data and multiple observers), it found that the public exposure doctrine did not apply.[174] From here, it went on to conclude "that individuals have a reasonable expectation of privacy in the whole of their physical movements," and that the government violated this expectation when it acquired Carpenter's CSLI data.[175]

In holding that individuals have a reasonable expectation of privacy in the whole of their physical movements, *Carpenter* has been widely described as a radical change in the Court's Fourth Amendment jurisprudence. For example, Orin Kerr states that *Carpenter* creates an entirely new type of search:

> *Carpenter* holds, for the first time, that a search occurred without it being a taking of information from any particular place, thing, or person. . . . [T]he government simply ended up with too much information about someone. How it ended up with too much information isn't particularly relevant in the Court's view. The point is the result, not the process.[176]

169. U.S. Dep't of Justice v. Reporters Comm. for Freedom of the Press, 489 U.S. 749, 749–50 (1989).

170. *Carpenter*, 138 S. Ct. at 2211.

171. *Id.* at 2208.

172. *Id.* at 2217–19.

173. *Id.*

174. *See* Evan Caminker, *Location Tracking and Digital Data: Can Carpenter Build a Stable Privacy Doctrine?*, 2018 SUP. CT. REV. 411, 437 (2018) (discussing this change in the framing of the question).

175. *Carpenter*, 138 S. Ct. at 2217, 2219 (citations omitted).

176. Orin Kerr, *When Does a* Carpenter *Search Start—and When Does It Stop?*, LAWFARE (July 6, 2018, 10:24 AM) (emphasis omitted), https://www.lawfareblog.com/when-does-carpenter-search-start-and-when-does-it-stop [https://perma.cc/TJ4Q-KJVJ] [hereinafter Kerr, *When Does a*

A similar point is made by others who suggest that *Carpenter* adopts the "mosaic theory" of privacy violations.[177] The mosaic theory has been articulated in different ways, but the core idea is that the aggregation of information about a person can violate a reasonable expectation of privacy because it provides a picture that is greater than the sum of its parts.[178] For example, in Justice Sotomayor's *Jones* concurrence, she concludes that aggregating publicly exposed location information could violate a reasonable expectation of privacy because it could reveal a person's "familial, political, professional, religious, and sexual associations."[179] This language is, notably, reiterated by the majority in *Carpenter*.[180]

If it were true that *Carpenter* recognized a new type of privacy violation based entirely on "the result, not the process,"[181] the case would pose a challenge to my claim that a core difference between privacy violations and losses is that violations are path-based whereas losses are outcome-based. But this interpretation of the case conflates two different issues. While the *Carpenter* Court quotes Sotomayor's language in *Jones*, it does not follow her in concluding that the mere aggregation of publicly exposed location data can violate a reasonable expectation of privacy.[182] To see the error underlying this widespread reading of the case, it is necessary to differentiate between two different questions that arise in the case—questions that track the privacy loss/violation distinction.

The first question is whether the public exposure doctrine excludes Carpenter's CSLI data from the scope of the Fourth Amendment's privacy protections. Note that this is a descriptive privacy *loss* question, asking whether Carpenter has already lost privacy in his physical movements. And it is in the Court's answer to this question—not the normative privacy *violation*

Carpenter *Search Start—and When Does It Stop?*]; *see also* ORIN S. KERR, THE DIGITAL FOURTH AMENDMENT: IMPLEMENTING CARPENTER (forthcoming 2020) (manuscript at 6), *available at* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3301257 [https://perma.cc/XE5A-FFHC] (describing the new test created by *Carpenter*).

177. *See, e.g.*, Caminker, *supra* note 174, at 436–43 (discussing "[t]he Court's first-ever embrace of a mosaic-defined search").

178. For example, as articulated by Stephanie Foster:

> The mosaic theory asserts that individually meaningless pieces of information, when aggregated, combine to create a revealing "mosaic," which is far more intrusive than any one piece of information. When viewed all together, the intimate picture painted by the mosaic violates an individual's reasonable expectation of privacy and therefore constitutes a search under the Fourth Amendment.

Stephanie Foster, *Should the Use of Automated License Plate Readers Constitute a Search After* Carpenter v. United States*?*, 97 WASH. U. L. REV. 221, 233 (2019) (citation omitted).

179. United States v. Jones, 565 U.S. 400, 415 (2012) (Sotomayor, J., concurring).

180. *Carpenter*, 138 S. Ct. at 2217.

181. Kerr, *When Does a* Carpenter *Search Start—and When Does It Stop?*, *supra* note 176.

182. *See Carpenter*, 138 S. Ct. at 2223.

question—that it could be said to adopt a version of the mosaic theory.[183] As explained above, the Court holds that the public exposure doctrine does not apply because "the whole" of Carpenter's physical movements was never exposed.[184] For this reason, this information is eligible for Fourth Amendment protection, giving rise to the second question.

The second question is whether the government's acquisition of CSLI data violated a reasonable expectation of privacy protected by the Fourth Amendment. On this question, the Court does not adopt an aggregation or mosaic based theory, but rather recognizes a violation based on the *means* by which the government accessed this information. The Court's decision is based on the fact that tracking via CSLI data allows "tireless and absolute surveillance" that is "retrospective," "nearly infallible," and entails "practically no expense."[185] The Court explains:

> Whether the Government employs its own surveillance technology as in *Jones* or leverages the technology of a wireless carrier, we hold that an individual maintains a legitimate expectation of privacy in the record of his physical movements *as captured through CSLI*. The location information obtained from Carpenter's wireless carriers was the product of a search.[186]

The fact that the means of access was central to the holding is confirmed by the Court's subsequent clarification of its scope: "Our decision today is a narrow one. We do not . . . call into question conventional surveillance techniques and tools, such as security cameras."[187] If the government had accessed the same information via a different permissible means, there would have been no violation.

Thus, *Carpenter* does not recognize that data aggregation itself violates a privacy right. It is not the case that the government merely ended up with too much information. Rather, the violation was based on the technology that provided access to the aggregated data. *Carpenter* thus supports my theory of the difference between privacy losses and violations, which in turn clarifies what is revolutionary about *Carpenter*—and what is not. Attention to the loss/violation distinction clarifies that the Court did not establish an entirely

183. At times, Caminker suggests that the Court adopted the mosaic theory in this way; but at others, he suggests that it adopted the mosaic theory as a theory of privacy violations. *Compare* Caminker, *supra* note 174, at 437 (discussing the mosaic theory in relation to the public exposure question), *with id.* at 439 (discussing "[t]he Court's first-ever embrace of a mosaic-defined search").

184. *See supra* notes 172–76 and accompanying text.

185. *Carpenter*, 138 S. Ct. at 2218–19; *see also* Susan Freiwald & Stephen Wm. Smith, *The* Carpenter *Chronicle: A Near-Perfect Surveillance*, 132 HARV. L. REV. 205, 219–22 (2018) (discussing the Court's multifactor analysis).

186. *Carpenter*, 138 S. Ct. at 2217 (emphasis added).

187. *Id.* at 2220.

new type of search that occurs through mere aggregation, but rather recognized the possibility of privacy in public.

The same lessons can be drawn from *Reporters Committee*, which has surprising parallels to *Carpenter*, but in the context of the government's *disclosure* of aggregated data under the Freedom of Information Act ("FOIA") (rather than government's acquisition of data under the Fourth Amendment).[188] The question in *Reporters Committee* was whether FOIA's privacy exemption, which restricts disclosures that "could reasonably be expected to constitute an unwarranted invasion of personal privacy,"[189] applied to FBI "rap sheets" that aggregated criminal information about individuals from various public sources.[190] The Court held that the disclosure of this information would constitute an unwarranted invasion of personal privacy,[191] and on this basis, scholars have cited *Reporters Committee* as the first Supreme Court case recognizing that data aggregation can violate a privacy right.[192] As with the literature on *Carpenter*, however, this conclusion is mistaken because it fails to differentiate between two questions that track the loss/violation distinction.

The first question (the privacy loss question) is whether the aggregated data can be considered private if the underlying data is in the public domain. The plaintiffs in *Reporters Committee* advanced an argument similar to the government's "public exposure" argument in *Carpenter*,[193] and the Court rejected it on similar grounds, concluding that the aggregated data provided a picture of the individuals that was not in fact public.[194] This conclusion has, moreover, been read as recognizing a privacy violation via data aggregation.[195] But as with *Carpenter*, this reading is mistaken.

When the Court in *Reporters Committee* addressed the second question (the privacy violation question), it explicitly rejected the claim that data aggregation violated a privacy right, stating that the Constitution "does not prohibit such a compilation."[196] Instead, it held that that *disclosure* of the

188. U.S. Dep't of Justice v. Reporters Comm. for Freedom of the Press, 489 U.S. 749, 751 (1989).

189. *Id.* (quoting 5 U.S.C. § 552(b)(7)(C) (2012)).

190. *Id.*

191. *Id.* at 780.

192. *See, e.g.*, Solove, *A Taxonomy of Privacy*, *supra* note 1, at 509.

193. *Reporters Comm.*, 489 U.S. at 757 (explaining that request was justified on the grounds that "[the rap sheet] contained 'matters of public record'").

194. *Id.* at 764 ("[T]here is a vast difference between the public records that might be found after a diligent search of courthouse files, county archives, and local police stations throughout the country and a computerized summary located in a single clearinghouse of information.").

195. *See, e.g.*, Solove, *A Taxonomy of Privacy*, *supra* note 1, at 509.

196. *Reporters Comm.*, 489 U.S. at 770. The Court acknowledged that the accumulation of personal information posed a "threat to privacy." *Id.* (quoting Whalen v. Roe, 429 U.S. 589, 605 (1977)). But it is important to differentiate between threats to privacy and their materialization for reasons discussed *supra* Section IV.A.1.

aggregated data *by the government* could cause a privacy violation. Thus, *Reporters Committee* recognized a path-based privacy right—a right to prevent the flow of information from the government to another party. Furthermore, its rationale for recognizing this right was also path-based; it was because the government was able to require access to the data (i.e., a means of access not available to the public) that it had a duty of non-disclosure.[197] So again, the case supports—and is clarified by—the theory of privacy that I have proposed.

Finally, as a normative matter, the fact that the Court did not recognize a privacy right against data aggregation in either of these cases should be seen as a good thing. There are two different sets of reasons for this.

The first arises from the fact that a right against data aggregation would impose an outcome-based restriction on data gathering, rather than a path-based restriction. This feature of the right is problematic in various ways that have been explored in depth in the critical literature on the mosaic theory. Some of the problems identified in this literature are specific to the Fourth Amendment context, but the core problems apply more generally. For example, Orin Kerr has identified a dizzying set of questions that judges would need to answer in order to enforce such a right,[198] leading many to conclude that the theory is unworkable.[199] These include hard categorization questions, such as whether data collected through different forms of human and technological observation should be grouped together or separated in conducting the mosaic analysis—and whether this should change if different people collected the data for different purposes. In addition, there are hard quantification questions, such as whether a technology that records the location of a person at 12 PM every day for five days should be seen as providing five seconds or five days of location data. To answer these and related questions, it seems that courts would need to draw unprincipled lines.[200]

In addition, and perhaps even more troubling, an outcome-based restriction would pose problems for those who want to conform their conduct to the law *ex ante*. For example, a person who collects data using different tools at different times might often have no way of knowing whether he is going to end up with a privacy-violating mosaic. Further, data gathering that is legal at the time it is conducted could, at any point in the future, retroactively become unlawful if the data subsequently becomes part of a privacy-violating mosaic. For these reasons, amongst others identified in the literature, the mosaic theory is problematic as a theory of privacy violations (though not as a theory

197. *Reporters Comm.*, 489 U.S. at 770.

198. *See generally* Orin S. Kerr, *The Mosaic Theory of the Fourth Amendment*, 111 MICH. L. REV. 311 (2012) [hereinafter Kerr, *The Mosaic Theory of the Fourth Amendment*] (identifying a list of questions that judges must address in order to enforce an outcome-based restriction).

199. *See* Rebecca Lipman, *Protecting Privacy with Fourth Amendment Use Restrictions*, 25 GEO. MASON L. REV. 412, 453 (2018).

200. *See* Kerr, *The Mosaic Theory of the Fourth Amendment*, *supra* note 198, at 346–47.

of privacy losses). This is one set of reasons why privacy rights should—as a normative matter—be defined as imposing path-based restrictions, not outcome-based restrictions.[201]

The second set of reasons to be concerned about a right against mere data aggregation—which is also highlighted by the loss/violation distinction that I have proposed—arises from a problem with the standard explanation of how data aggregation violates privacy. According to this account, aggregation violates privacy by revealing sensitive information that was not visible in the disaggregated data. For example, as articulated by Solove: "People expect certain limits on what is known about them and on what others will find out. Aggregation upsets these expectations, because it involves the combination of data in new, potentially unanticipated ways to reveal facts about a person that are not readily known."[202] A similar rationale is offered by Justice Sotomayor in *Jones*, where she states that reasonable expectations of privacy are violated when personal location information about individuals is "recorded and aggregated in a manner that enables the Government to ascertain, more or less at will, their political and religious beliefs, sexual habits, and so on."[203]

Paying attention to the path-based element of privacy violations reveals a potential problem with this account. For example, in Sotomayor's example, it highlights that the government is not actually acquiring political/religious/sexual information about a person. Rather, the government is inferring it—and the fact that information is being inferred, rather than discovered, is relevant in ways that have gone unexplored. On one hand, it could be argued, this account of how aggregation violates privacy rights avoids some of the problems identified above, as it locates the violation in the decision to analyze and draw inferences from aggregated data (rather than the mere fact that aggregation has occurred). On the other hand, the claim that aggregation can violate privacy on these grounds entails the premise that an inference is a means of access that can violate privacy rights. This premise, which has not been recognized or defended in the literature, is problematic for reasons that I will explore in Section V.C below. Before doing so, however, it will be helpful to first clarify the issue of unconsented use.

2. No Right Against Unconsented Use

The second possible challenge to my claim that privacy rights impose path-based restrictions on access comes from the view that the mere

201. This point has been recognized and developed by Gray and Citron, who argue that "quantitative privacy" should not be defined in terms of "how much information is gathered in a particular case" but rather on "how information is gathered." David Gray & Danielle Citron, *The Right to Quantitative Privacy*, 98 MINN. L. REV. 62, 71 (2013) (emphasis omitted).

202. Solove, *A Taxonomy of Privacy*, *supra* note 1, at 508.

203. United States v. Jones, 565 U.S. 400, 416 (2012) (Sotomayor, J., concurring).

unconsented use of personal data can violate privacy rights.[204] If the best theory of privacy rights includes rights that restrict unconsented use per se —thereby protecting pure informational autonomy—these rights would pose a challenge to my account. But as with aggregation, this expansive interpretation of use restrictions should be rejected.

To start with matters of existing law, this position is at best only plausible with respect to a limited set of privacy rights, and even in the case of this limited set, the claim should ultimately be rejected. There are four core points here: two about limits, and two about deeper problems.

First, none of the foundational areas of privacy law, including the Fourth Amendment,[205] the common law privacy torts,[206] and the constitutional right to informational privacy,[207] protect a right to restrict or control the use of one's information. Insofar as there is such a right, it is to be found in sector-specific privacy laws that have their origin in the set of "principles of fair information practice" identified in a 1973 report by the U.S. Department of Health, Education, and Welfare. Among these principles is a "purpose specification" principle restricting use: "There must be a way for an individual to prevent information about him obtained for one purpose from *being used* or *made available* for other purposes without his consent."[208] Whether this is accurately described as a principle of privacy, as is widely suggested, is a question to which I will return below.

Second, even within the limited set of laws that adopt the purpose specification principle, use itself is not always restricted. For example, the Gramm–Leach–Bliley Act of 1999 has a section titled "[l]imits on reuse of information," but this section actually only restricts access, limiting the parties

204. *See, e.g.*, Solove, *Conceptualizing Privacy*, *supra* note 3, at 1108 ("Privacy . . . involves more than avoiding disclosure; it also involves the individual's ability to ensure that personal information is used for the purposes she desires."); *see also* Solove, *A Taxonomy of Privacy*, *supra* note 1, at 518–20 (developing this conception of privacy).

205. *See* Elizabeth E. Joh, *Policing by Numbers: Big Data and the Fourth Amendment*, 89 WASH. L. REV. 35, 63 (2014) ("The Fourth Amendment is primarily interested in the legitimacy of *how* information is acquired. If the acquisition is permissible, how the police *use* that information thereafter is generally not subject to an additional Fourth Amendment challenge." (second emphasis added) (citation omitted)). While Ric Simmons suggests that "we have seen growing numbers of lower courts turning to use restrictions to solve some of the modern problems posed by technology and the Fourth Amendment[,]" he acknowledges that these are not "use restrictions" in the traditional senses. Ric Simmons, *The Mirage of Use Restrictions*, 96 N.C. L. REV. 133, 180, 198 (2017). In fact, these restrictions are better characterized as a type of access restriction, as the courts have held that when the police seize data in bulk (e.g., through seizing a hard drive), the act of accessing the data can constitute a new search that must be independently justified. *Id.* at 136.

206. Anupam Chander, *How Law Made Silicon Valley*, 63 EMORY L.J. 639, 664–65 (2014); Paul M. Schwartz, *Privacy and Democracy in Cyberspace*, 52 VAND. L. REV. 1609, 1634 (1999).

207. Daniel J. Solove, *Privacy and Power: Computer Databases and Metaphors for Information Privacy*, 53 STAN. L. REV. 1393, 1436–37 (2001).

208. U.S. DEP'T OF HEALTH, EDUC., & WELFARE, RECORDS, COMPUTERS, AND THE RIGHTS OF CITIZENS 41 (1973) (emphasis added).

to whom personal information may be disclosed.[209] Furthermore, many statutes that do restrict use connect this restriction to an access restriction.[210] There are only a few statutes that impose access-independent use restrictions.[211]

Third, while it has been suggested that these use restrictions provide rights against unconsented use, embodying the idea of privacy-as-control,[212] this characterization misses an important point: that the use and transfer of the data is often not under the person's control.[213] Although unconsented uses beyond those authorized by the statutes are not permitted, the data subject is not in control of the uses that are authorized. For example, under the Fair Credit Reporting Act, consumers are not provided with the opportunity to opt-out of the creation, disclosure, and use of credit reports about them,[214] which suggests that the use restrictions are really about protecting fairness.

Fourth, the claim that these statutes provide rights against unconsented use is undercut by the fact that the restrictions are eliminated by anonymization (including in limited forms, such as pseudonymization and data perturbation). This is the case with the Freedom of Information Act,[215]

209. 15 U.S.C. § 6802(c) (2012). Likewise, the Video Privacy Protection Act of 1988 only imposes restrictions on the parties to whom records may be disclosed, and the Cable Communications Policy Act of 1984 only restricts data collection and storage. 47 U.S.C. § 551 (2012).

210. For example, under the Fair Credit Reporting Act, a consumer reporting agency can provide a credit report to a third party only for limited purposes (an access restriction), and any subsequent use must be in accordance with one of these permissible purposes (the use restriction). *See* 15 U.S.C. §§ 1681b(a), 1681e(e). The same structure can be found in the Driver's Privacy Protection Act. *See* 18 U.S.C. § 2721 (2012) (the access restriction); *id.* § 2724 (the use restriction).

211. For example, the Health Insurance Portability and Accountability Act regulations impose a wide range of restrictions on the use of medical information (beyond those necessary for treatment, payment, and health care operations), 45 C.F.R. § 164.508 (2002), and the Federal Election Campaign Act states that documents filed with the Commission "may not be sold or used by any person for the purpose of soliciting contributions or for commercial purposes." 52 U.S.C. § 30111 (2012).

212. Michael D. Birnhack, *A Quest for a Theory of Privacy: Context and Control*, 51 JURIMETRICS 447, 449 (2011) (reviewing NISSENBAUM, *supra* note 13).

213. *See generally* Paul M. Schwartz, *Beyond Lessig's* Code *for Internet Privacy: Cyberspace Filters, Privacy Control, and Fair Information Practices*, 2000 WIS. L. REV. 743 [hereinafter Schwartz, *Beyond Lessig's* Code *for Internet Privacy*] (stating that the FIPP framework is an alternative to privacy-as-control).

214. Janet Dean Gertz, Comment, *The Purloined Personality: Consumer Profiling in Financial Services*, 39 SAN DIEGO L. REV. 943, 980 (2002).

215. Dep't of the Air Force v. Rose, 425 U.S. 352, 381 (1976); ACLU v. Dep't of Def., 543 F.3d 59, 85 (2d Cir. 2008), *vacated*, 130 S. Ct. 777 (2009).

Privacy Act of 1974,[216] Health Insurance Portability and Accountability Act,[217] Fair Credit Reporting Act,[218] Video Privacy Protection Act,[219] and many state genetic privacy laws.[220] Further, this is not only a feature of these specific statutes, but rather a core part of the underlying FIPP framework. As Barocas and Nissenbaum explain, the framework provides two options: consent or anonymize.[221] Thus, this approach only provides limited rights of control when the information is connected to the individual—when use would entail access. It does not provide a right against mere unconsented use.

Further, as a normative matter, there are good reasons why privacy law should not include rights that protect against mere unconsented use, independent of the way in which it is used (e.g., to create new forms of access). To do so would be to reinforce the mistaken conflation of autonomy and privacy. The conflation of these issues dates back at least as far as the Supreme Court cases describing constitutionally protected rights to contraception and abortion as "privacy rights."[222] This categorization of these rights was widely criticized as creating conceptual confusion,[223] and the Supreme Court has recently remedied this error. It now characterizes these reproductive rights (and other related rights) in terms of autonomy rather than privacy.[224] Privacy law should not reintroduce the confusion.

To state that privacy and autonomy should not be conflated is not to deny there is a connection between them. For example, it is clear that autonomy interests provide one strong justification for granting privacy rights. As discussed in Part II, autonomy is at the core of many accounts of why privacy rights are important.[225] It is a mistake, however, to then interpret these privacy rights as protecting a broader set of informational autonomy interests. This reasoning is similar to a common logical error, the "fallacy of the converse," in which one starts with a true proposition and then invalidly infers its

216. *See* Privacy Act of 1974, 5 U.S.C. § 552a(b)(1)–(12) (2012).

217. Lawrence O. Gostin & James G. Hodge, Jr., *Personal Privacy and Common Goods: A Framework for Balancing Under the National Health Information Privacy Rule*, 86 MINN. L. REV. 1439, 1458–59, 1459 n.73 (2002).

218. Benjamin Charkow, Note, *The Control over the De-Identification of Data*, 21 CARDOZO ARTS & ENT. L.J. 195, 208–09 (2003).

219. *In re* Nickelodeon Consumer Privacy Litig., 827 F.3d 262, 279 (3d Cir. 2016).

220. *See, e.g.*, N.J. STAT. ANN. §§ 10:5-43 to 10:5-49 (West 2013); OR. REV. STAT. ANN. §§ 192.531–192.549 (West 2007).

221. Solon Barocas & Helen Nissenbaum, *Computing Ethics: Big Data's End Run Around Procedural Privacy Protections*, 57 COMM. ACM 31, 31 (2014).

222. *See* Roe v. Wade, 410 U.S. 113, 152–54 (1973); Griswold v. Connecticut, 381 U.S. 479, 482–86 (1965).

223. *See, e.g.*, TRIBE, *supra* note 11, at 1352; Gavison, *supra* note 10, at 438–39; Gross, *supra* note 6, at 38; Henkin, *supra* note 11, at 1410–11.

224. *See generally* Greene, *supra* note 45 (describing the Supreme Court's shift from privacy to liberty/autonomy as the constitutional basis for the freedom to make fundamental life decisions).

225. *See supra* Section II.A.1.

converse. In this case (stated simply), "protecting autonomy requires protecting privacy" is mistakenly taken to imply "protecting privacy requires protecting autonomy." Thus, while autonomy interests justify privacy rights, this does not mean that privacy rights should include the protection of informational autonomy.[226]

In arguing that the unconsented use of personal information should not be restricted by privacy rights, I am not arguing that it should not be restricted on other grounds. For example, while the Fair Information Practice Principles have often been classified as an aspect of privacy law,[227] it seems to me that they are actually (as the name suggests) about fairness. This is not, however, merely a linguistic point. Classification here has normative significance. When the interests implicated by unconsented use are properly identified, we may find that they do not justify rights as expansive as privacy rights.

C. INFERENCES OF PERSONAL INFORMATION

Having argued that the aggregation and unconsented use of personal information do not violate privacy rights themselves (i.e., independently of violating an access restriction), I will now turn to a final question that emerges from my analysis, which is whether inferences can violate privacy rights.

It is clear that the analysis of disclosed personal data can reveal personal facts that were not knowingly disclosed. Perhaps the most often-cited example of this is the case of Target correctly inferring the early-stage pregnancy of customers based on their purchasing pattern of items that were not clearly linked to pregnancy, such as unscented lotion.[228] This example will soon seem quaint, however, as machine-learning algorithms infer significantly more complex personal traits from seemingly irrelevant data collected across disparate domains of life.

It is often said that these types of discoveries violate privacy rights,[229] but this conclusion relies on the assumption that inferences can do so—an assumption that, as far as I can tell, no one has explored in depth. Perhaps the fact that inferences can clearly cause significant privacy losses, and in ways that are often not reasonably foreseeable, makes it seem equally clear that they can violate privacy rights. As my analysis has highlighted, however, what privacy rights protect is not a reasonable expectation that privacy will not be

226. *Cf.* Schwartz, *Beyond Lessig's* Code *for Internet Privacy*, *supra* note 213, at 759 ("Protection of the capacity for self-determination requires a setting of limits on the collection of personal data, but it does not call for privacy-control as a central means of achieving these limits.").

227. *See, e.g.*, DANIEL J. SOLOVE & PAUL M. SCHWARTZ, INFORMATION PRIVACY LAW 36–39 (6th ed. 2018) (including FIPP rules in an overview of statutory privacy law).

228. *See, e.g.*, Crawford & Schultz, *supra* note 1, at 94–95.

229. *See, e.g.*, *id.* at 96–109; Woodrow Hartzog et al., *Inefficiently Automated Law Enforcement*, 2015 MICH. ST. L. REV. 1763, 1789; Sheri B. Pan, Note, *Get to Know Me: Protecting Privacy and Autonomy Under Big Data's Penetrating Gaze*, 30 HARV. J.L. & TECH. 239, 259 (2016); Benjamin Zhu, Note, *A Traditional Tort for a Modern Threat: Applying Intrusion upon Seclusion to Dataveillance Observations*, 89 N.Y.U. L. REV. 2381, 2382–87 (2014).

lost; rather, they protect reasonable expectations that privacy will not be lost in certain ways. Thus, the core unexplored questions here are whether privacy rights currently restrict inferences and whether they should.

My argument will be that inferences do not violate established privacy rights in the ways that have been suggested, and furthermore, that the creation of privacy rights that could be violated by inferences would raise significant normative concerns that may, at the very least, justify restricting such rights in ways that have received insufficient attention.

1. Fourth Amendment Confusion

It is often stated in passing that the Supreme Court's decision in *Kyllo v. United States* held that the Fourth Amendment can be violated by either (a) the inference of personal information in which one has a reasonable expectation of privacy, or (b) the acquisition of data that allows for such inferences.[230] In fact, the Court reached the opposite conclusion in this landmark technology case, which appears to be the only case of privacy law that explicitly addresses the status of inferences in privacy violations.[231] The Court's statement that an inference cannot "insulate[] a search," which it reiterated without discussion in *Carpenter v. United States*,[232] has been the source of much confusion. In order to understand what the *Kyllo* Court meant by this, and the status of inferences more generally, it is necessary to clarify

230. *See, e.g.*, Leslie A. Lunney, *Has the Fourth Amendment Gone to the Dogs?: Unreasonable Expansion of Canine Sniff Doctrine to Include Sniffs of the Home*, 88 OR. L. REV. 829, 855 (2009) ("[T]he thermal scan was a 'search' because it made technology-assisted inferencing about the interior of a home possible."); *Leading Cases: Constitutional Law*, 115 HARV. L. REV. 306, 349 n.34 (2001) ("The Court also dispensed with the dissent's suggestion that information learned through an inference could not be a search."); Matthew J. Tokson, *The Content/Envelope Distinction in Internet Law*, 50 WM. & MARY L. REV. 2105, 2146–47 (stating that under *Kyllo* and *Karo*, the collection of public information that allows for the inference of private information can constitute a search); Christopher M. Pardo, *Driving off the Face of the Fourth Amendment: Weighing* Caballes *Under the Proposed "Vehicular Frisk" Standard*, 43 VAL. U. L. REV. 113, 140 n.131 (2008) ("*Kyllo* establishes both that the use of sensory enhancing devices and inferences drawn from them are searches pursuant to the Fourth Amendment."); Sean D. Thueson, *Fourth Amendment Search—Fuzzy Shades of Gray: The New "Bright-Line" Rule in Determining When the Use of Technology Constitutes a Search.* Kyllo v. United States, *121 S. Ct. 2038 (2001).*, 2 WYO. L. REV. 169, 201 (2002) ("The rule in *Kyllo* is too broad because inferences made by the police can become illegal searches.").

231. There are a few cases, however, that discuss whether the ability to draw an inference can *eliminate* a reasonable expectation of privacy. *Compare* Walter v. United States, 447 U.S. 649, 650 (1980) (holding that the defendants had a reasonable expectation of privacy in pornographic films, even though their boxes had allowed by government to "draw inferences about what was on the films"), *with* Arkansas v. Sanders, 442 U.S. 753, 764 n.13 (1979) (holding that "some containers (for example a kit of burglar tools or a gun case) by their very nature cannot support any reasonable expectation of privacy because their contents can be inferred from their outward appearance").

232. Carpenter v. United States, 138 S. Ct. 2206, 2218 (2018) (arguing that although CSLI did not provide a precise location, this did not mean that government access to the data should be permitted because "the Court has already rejected the proposition that 'inference insulates a search.'" (quoting Kyllo v. United States, 533 U.S. 27, 36 (2001))).

some of the significant complexities of the case. Although the case is nominally about thermal-imaging, which might seem low-tech, the Court addressed complex epistemological issues that are directly relevant to the machine-learning algorithms of today.

The core question in the case was whether the use of a thermal-imaging device to detect heat within a home violated a reasonable expectation of privacy protected by the Fourth Amendment.[233] The Court held that it did, but the case was decided in a five to four split,[234] with the disagreement focusing on the role of inferences in the case and in the law. This disagreement started with the dissent's claim that the Court ignored "a distinction of constitutional magnitude between 'through-the-wall surveillance' that gives the observer or listener *direct access* to information in a private area" and "'off-the-wall' surveillance" that allows one "to draw *inferences* from information in the public domain."[235] According to the dissent, the Court's finding of a search established "[f]or the first time in its history . . . that an inference can amount to a Fourth Amendment violation."[236] In response, the majority accused the dissent of adopting "the novel proposition that inference insulates a search," which it stated "is blatantly contrary to" well-established precedent.[237] From these statements, it might seem that the majority and dissent reached opposing conclusions about the legal status of inferences. But in fact, they did not. In order to understand their positions —and what the Court actually held—it is necessary to unpack several issues.

First, it is important to recognize that while the dissent seems to attach legal significance to the distinction between direct "through-the-wall" and indirect "off-the-wall" data gathering—which could be relevant to the operation of many algorithms (which are arguably "off the wall")—it cannot be the case that the distinction has constitutional relevance in and of itself. As the majority notes, this position would be incompatible with clear cases of privacy violations: for example, the impermissibility of using directional microphones to listen to conversations inside a home, even though they measure sound emanating from off the house.[238] In addition, this position would be incompatible with the dissent's view that a more sophisticated thermal-imaging device that reveals activities in the home would be

233. *Kyllo*, 533 U.S. at 29.

234. *See id.* The opinion of the Court was written by Justice Scalia and joined by Justices Souter, Thomas, Ginsburg, and Breyer. The dissent was written by Justice Stevens and joined by Chief Justice Rehnquist and Justices O'Connor and Kennedy.

235. *Id.* at 41 (Stevens, J., dissenting) (emphasis added).

236. *Id.* at 44 (citation omitted); *see also id.* at 49 ("[T]he Court effectively treats the mental process of analyzing data obtained from external sources as the equivalent of a physical intrusion into the home. As I have explained, however, the process of drawing inferences from data in the public domain should not be characterized as a search." (citation omitted)).

237. *Id.* at 36 (majority opinion).

238. *Id.* at 35–36.

impermissible, even though such a device would also only measure heat "off-the-wall" rather than "through-the-wall."

For these reasons, the *direct/indirect* distinction must be relevant only to the extent that it maps onto the *access/inference* distinction that the dissent also highlights. This is a core issue that the government raised at oral argument, when the Deputy Solicitor General highlighted that police could not learn "directly from the imager" that heat was being generated inside the house. Rather, he explained, the police had to infer this, as it was possible that the heat had been produced in some other way. For example, the walls could have been "unduly heated up by the sun."[239] This "reduce[d] the specificity and directness, the linearality of any inference" that could be drawn from the heat sensor. [240] Likewise, the dissent highlighted that "the only conclusions the officers reached concerning the interior of the home were at least as *indirect* as those that might have been *inferred* from the contents of discarded garbage, . . . or pen register data, . . . or, as in this case, subpoenaed utility records."[241] This feature of the technology—the fact that inferences were required to gain knowledge of the inside of the home—provides the best explanation of why the dissent concluded that the technology "did not obtain 'any information regarding the interior of the home.'"[242]

In drawing this conclusion, however, the dissent failed to recognize the difference between two different questions, which is worth highlighting because it will be relevant to many technologies. The first is the question of whether one knows that one has accessed a piece of information. The second is the question of whether one has accessed it. In this case, the police might not have known with certainty that the technology was revealing heat inside the house. Because of the potentially confounding factors, any conclusion about the inside of the house was an uncertain inference. But this does not mean that the technology was not in fact measuring heat from inside the house. One is a question of what the police knew; the other is a question of what they did.

Unfortunately, neither side recognized this crucial distinction, which has resulted in significant confusion about key matters of law. This started with the majority thinking that the dissent was making the "extraordinary assertion that anything learned through 'an inference' cannot be a search," to which it replied:

239. Transcript of Oral Argument at 41–42, Kyllo v. United States, 533 U.S. 27 (2001) (No. 99-8508).

240. *Id.* at 42; *see also id.* ("There isn't a one-to-one correspondence between heat on the exterior of the structure and heat on the interior of the structure.").

241. *Kyllo*, 533 U.S. at 44 (Stevens, J., dissenting) (emphasis added) (citations omitted).

242. *Id.* (emphasis omitted) (quoting the majority opinion). While the dissent does not spell this out explicitly, this seems to be the best explanation of this conclusion. This feature also underlies the dissent's comparison of the technology with other practices that were permissible under well-established precedent.

> [T]he novel proposition that inference insulates a search is blatantly contrary to *United States v. Karo*, where the police "inferred" from the activation of a beeper that a certain can of ether was in the home. The police activity was held to be a search, and the search was held unlawful.[243]

But as the dissent clarified, this is not what it was saying:

> Although the Court credits us with the "novel proposition that inference insulates a search," our point simply is that an inference cannot be a search, contrary to the Court's reasoning. Thus, the Court's use of *United States v. Karo*[] to refute a point we do not make underscores the fact that the Court has no real answer (either in logic or in law) to the point we do make. Of course, *Karo* itself does not provide any support for the Court's view that inferences can amount to unconstitutional searches.[244]

In order to unpack the disagreement here, it is necessary to first clarify a point about *Karo*. While the police in this case did infer that the can of ether was in the house, as the majority's comment highlights, *Karo* did not hold that this inference constituted the search. Rather, it held that the "monitoring of a beeper in a private residence" (i.e., the acquisition of the raw data underlying the inference) violated the Fourth Amendment.[245] Thus, the *Kyllo* majority's discussion of inferences in *Karo* was *not* meant to establish that an inference can constitute a search, but rather that the acquisition of the raw data can constitute a search even if inferences are required to interpret it.

Unfortunately, the majority's description of the holding in *Karo* did not only cause confusion in the dissent, but also in the literature on the legal status of inferences. It has led many to conclude that an inference can constitute a search and turn the underlying collection of data into a search. For example, Leslie Lunney concludes that the thermal scan in *Kyllo* was a search "because it made technology-assisted inferencing about the interior of a home possible,"[246] and this interpretation of the majority's statement is widespread.[247] However, as should now be clear, this is not what the majority meant. The majority further clarified this point in a footnote about its "insulate a search" comment:

> The dissent asserts that we have misunderstood its point, which is not that inference *insulates* a search, but that inference alone is *not* a search. If we misunderstood the point, it was only in a good-faith effort to render the point germane to the case at hand. The issue in

243. *Id.* at 36–37 (majority opinion) (citation omitted).

244. *Id.* at 44 n.3 (Stevens, J., dissenting) (emphasis omitted) (citations omitted).

245. United States v. Karo, 468 U.S. 705, 714 (1984).

246. Lunney, *supra* note 230, at 855.

247. *See* sources cited *supra* note 225.

> this case is not the police's allegedly unlawful inferencing, but their allegedly unlawful thermal-imaging measurement of the emanations from a house. We say such measurement is a search; the dissent says it is not, because an inference is not a search. We took that to mean that, since the technologically enhanced emanations had to be the basis of inferences before anything inside the house could be known, the use of the emanations could not be a search. But the dissent certainly knows better than we what it intends. And if it means only that an inference is not a search, we certainly agree.[248]

Here, the Court's reference to what is "known" about the inside of the house points to the actual nature of the disagreement between the majority and dissent, which I touched on earlier. It is a disagreement about the level of epistemic warrant that is needed to cause a privacy violation. Whereas the dissent asks if the technology provides the police with *knowledge* about the inside of the house, the majority asks if it provides them with *data* about the inside the house.

Further, because they start with different questions without recognizing it, the majority misinterprets the minority's legal position and vice versa. First, the majority: because the majority focuses on *data* acquisition and concludes that the technology does provide data about the inside of the house (as a factual matter), it believes that the dissent's denial of a search relies on the assumption that an inference can *insulate* a search (as a matter of law). This is the only way the majority is able to understand how the dissent reaches the conclusion that there is no search. Second, the dissent: because the dissent focuses on *knowledge* and concludes that the technology does not provide knowledge of the inside of the home (as a factual matter), it believes that the majority's finding of a search relies on the assumption that an inference can *constitute* a search (as a matter of law). This is the only way the dissent is able to understand how the majority reaches the conclusion that there is search.

Thus, while it might appear that the majority and dissent reach different legal conclusions in the case because they disagree about matters of fact, it is actually because they disagree on a matter of law. They disagree about the epistemic status that is required for a search, and therefore ask different questions about the facts.

The underlying question of which side is right about the epistemic requirement is outside the scope of this paper, but I will conclude with one brief observation on this point, which parallels my argument about the "epistemic merit" requirement for privacy losses (in Section IV.B). For here too, in the context of privacy violations, it seems that a means of access does not need to provide knowledge to violate the Fourth Amendment, as the *Kyllo* dissent assumes—but that the dissent is right to think that there is some epistemic requirement that must be satisfied. Under *Karo*, at least, it seems

248. *Kyllo*, 533 U.S. at 37 n.4 (citation omitted).

that a means of access can only violate a reasonable expectation of privacy if it provides one with grounds for forming a true belief about the data at issue; the mere fact that data is acquired by a technology is insufficient.[249]

In sum, despite their disagreements and contrary to a widespread reading of the case, both the majority and dissent in *Kyllo* agree that an inference cannot violate a reasonable expectation of privacy under the Fourth Amendment. Further, neither opinion supports the claim that an inference of personal information from data can transform the underlying collection of that data into a Fourth Amendment violation.

2. Problems with Restricting Inferences

It is unclear whether courts applying other sources of privacy law, including the various sources that have adopted a version of the "reasonable expectation of privacy" test,[250] will reach the same conclusion as the Supreme Court in *Kyllo*. It is clear from the privacy literature, however, that many would argue that they should not. Whether the law should ever recognize a privacy right against inferences is a question that goes beyond the scope of this paper, but to conclude my analysis of inferences, I will briefly note three reasons to think that a right against inferences would need to be more limited than has been recognized.

First and most importantly, a privacy right that restricted one's ability to infer private facts about others would impose restrictions on purely mental activity. This would violate foundational principles of morality and law.[251] Even indirect means of mind control have been found unconstitutional.[252] Thus, if privacy law were to recognize a right against inferences, the scope of the right would clearly need to be limited to exclude mental inferences. Some

249. This epistemic requirement is implicit in the Court's decision in *Karo*, where the police not only tracked the can of ether when it was inside a home (as discussed above), but also when it was inside a locker in a warehouse. While the Court held that the tracking in the home was a search (as discussed above and in *Kyllo*), it held that the tracking in the warehouse was not. The reason was that "the beeper informed the agents only that the ether was somewhere in the warehouse; it did not identify the specific locker in which the ether was located." *Karo*, 468 U.S. at 720. The locker "was identified only when agents traversing the public parts of the facility found that the smell of ether was coming from a specific locker." *Id.* at 720–21. The monitoring of "the beeper revealed nothing about the contents of the locker . . . and hence was not a search of that locker." *Id.* at 720. Thus, although the beeper was in fact transmitting its location from inside the locker, it provided the police with no way to form a belief about its location inside the locker, and for this reason was insufficient to constitute a search.

250. These include the common law privacy torts, the Freedom of Information Act, the Privacy Act of 1974, the constitutional right of information privacy, and various evidentiary privileges. Lior Jacob Strahilevitz, *A Social Networks Theory of Privacy*, 72 U. CHI. L. REV. 919, 985–86 (2005).

251. Adam J. Kolber, *Two Views of First Amendment Thought Privacy*, 18 U. PA. J. CONST. L. 1381, 1384–86 (2016); Gabriel S. Mendlow, *Why Is It Wrong to Punish Thought?*, 127 YALE L.J. 2342, 2345 (2018).

252. Mendlow, *supra* note 251, at 2369–70.

distinction between human and non-human inferences would need to be drawn and justified.

Second, even if the right only restricted computer-assisted inferential analysis, it would impose limits on free inquiry, independently of any act or outcome. There are strong normative grounds, and potentially constitutional grounds, to be concerned about such restrictions.[253] Further, to analyze whether privacy interests outweigh interests in free inquiry, one would need to differentiate between two different types of inquiry: data gathering versus data analysis. Privacy interests might provide strong justifications for restricting data gathering, but many of these justifications (for example, those based in a conception of private spaces) do not apply to restrictions on data analysis.[254] Thus, in the case of data analysis, free inquiry interests might outweigh the privacy interest in preventing inferences—just as the First Amendment right to free speech can outweigh the privacy right against disclosure.

Third, the affirmative basis for recognizing a right against inferences is limited in ways that have been obscured by the failure to differentiate between the different types of interests at stake. Imagine, for example, that a company's HR department uses a sophisticated algorithm to predict health problems from non-health data in job applications. While the applicants in this case certainly have an interest in preventing the inferences, this is only partly an interest in preventing access to their health information (i.e., a matter of privacy). It is also—and arguably more so—an interest in preventing the use of this information in making a hiring decision (i.e., a matter of fairness). But to protect the latter type of interests, the law should grant rights that target this harmful conduct (for example, a right against health-based discrimination), not a right against inferences.[255]

To be clear, I am not saying that the law should never recognize a right against inferences. Rather, my point is that the recognition of such a right would raise many concerns that have not been explored—or even recognized—in the literature. By drawing attention to these issues, my account of privacy highlights essential lines for future enquiry.

253. *See generally* Dana Remus Irwin, *Freedom of Thought: The First Amendment and the Scientific Method*, 2005 WIS. L. REV. 1479 (discussing the "First Amendment protection of scientific experiment"); Natalie Ram, *Science as Speech*, 102 IOWA L. REV. 1187 (2017) (arguing that restrictions on scientific inquiry implicate the First Amendment).

254. Concerns about information gathering might also provide a better explanation of the concerns that people have articulated in terms of inferences. Cary Coglianese & David Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, 105 GEO. L.J. 1147, 1176 n.119 (2017) (suggesting that although algorithms allow new kinds of inferences, "the underlying privacy problems that others have flagged appear to have been raised more by the collection of big data than by the use of machine-learning algorithms").

255. Of course, there may be cases in which it is difficult to prevent the harmful conduct, in which case preventing the inferences might be a justifiable second-best solution. But in this case, the right against the inferences would not be properly classified as a privacy right.

VI. CONCLUSION

It is widely thought that the core problems posed by new technologies of personal data mining and analysis, as well as their solutions, can be explained in terms of privacy. There is also growing agreement that a unified theory of privacy is unattainable. This Article demonstrates that these are both mistaken conclusions that derive from the conflation of privacy losses and violations, and it develops a theory of privacy that untangles these misunderstood concepts at the core of privacy law. In clarifying the outcome-based criteria that define privacy losses and their relationship with the path-based criteria that define privacy violations, this theory provides value across two domains. First, regarding the coherence of the law, it demonstrates how a unified theory of privacy rights is possible despite significant disagreement about their content. Second, regarding the law's content, it reveals foundational distinctions that have gone unrecognized, challenging orthodox views about how privacy rights are violated by the aggregation, unconsented use, and inference of personal data. It is possible that these practices should be restricted on other grounds, but when the actual interests at stake are identified, it may become clear that they do not justify restrictions that are as expansive as those that have been envisioned. Thus, recognizing the difference between losses and violations reveals both the unity and the limits of privacy.