JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY

VOLUME XXX · NUMBER 3
May 2025

ARTICLES

DISCUSSION

OPEN ACCESS FREE FOR AUTHORS AND READERS ISSN 1559-3061

JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY

VOLUME XXX · NUMBER 3
May 2025

ARTICLES

DISCUSSION

JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY
http://www.jesp.org
ISSN 1559-3061

The *Journal of Ethics and Social Philosophy* (*JESP*) is a peer-reviewed online journal in moral, social, political, and legal philosophy. The journal is founded on the principle of publisher-funded open access. There are no publication fees for authors, and public access to articles is free of charge. Articles are typically published under the CREATIVE COMMONS ATTRIBUTION-NONCOMMERCIAL-NODERIVATIVES 4.0 license, though authors can request a different Creative Commons license if one is required for funding purposes. Funding for the journal has been made possible through the generous commitment of the Division of Arts and Humanities at New York University Abu Dhabi.

JESP aspires to be the leading venue for the best new work in the fields that it covers, and it is governed by a correspondingly high editorial standard. The journal welcomes submissions of articles in any of these and related fields of research. The journal is interested in work in the history of ethics that bears directly on topics of contemporary interest, but does not consider articles of purely historical interest. It is the view of the editors that the journal's high standard does not preclude publishing work that is critical in nature, provided that it is constructive, well argued, current, and of sufficiently general interest. *JESP* also endorses and abides by the Barcelona Principles for a Globally Inclusive Philosophy, which seek to address the structural inequality between native and nonnative English speakers in academic philosophy.

JESP publishes articles, discussion notes, and occasional symposia. Articles normally do not exceed 12,000 words (including notes and references). *JESP* sometimes publishes longer articles, but submissions over 12,000 words are evaluated according to a proportionally higher standard. Discussion notes, which need not engage with work that was published in *JESP*, should not exceed 3,000 words (including notes and references). *JESP* does not publish book reviews.

Papers are published in PDF format at https://www.jesp.org. All published papers receive a permanent DOI and are archived both internally and externally.

Journal of Ethics and Social Philosophy
VOL. 30, NO. 3 · MAY 2025
https://doi.org/10.26556/jesp.v30i3.4023

CLARIFYING *PARENS PATRIAE*

James G. Dwyer

LEGAL SYSTEMS treat a third of the human population as nonautonomous, incompetent to govern their own lives. This includes persons of any age with incapacitating mental conditions (intellectual disability, mental illness, dementia, etc.) and minors. Recent decades have seen challenges to the categorization by lawmakers of individuals as such, with theorists urging greater respect for the views of adults with lesser cognitive capacity, of adolescents, and even of young children.[1] Without articulating and defending here a position on the complex matters of where the law should draw lines between those entitled to make life-determining decisions for themselves (with or without support) and those who need some entity to make certain decisions on their behalf (with their input as appropriate) and of how the law should treat transitions from one to the other, this article aims to clarify the role of the state in the lives of persons in the latter category.[2] Certainly, this category includes newborns and prelingual infants, and arguably, it also includes older children in connection with relatively complex and momentous choices such as medical treatment and type of schooling. It also includes a significant percentage of adults, such as those with advanced dementia and those unable to make or communicate choices because of severe disability, mental illness, or brain injury.[3]

1 See Archard and Uniacke, "The Child's Right to a Voice"; and Arstein-Kerslake, *Restoring Voice to People with Cognitive Disabilities*. The United Nations Convention on the Rights of the Child (UNCRC) urges greater voice and freedom for children. See art. 12 ("State parties shall assure to a child who is capable of forming his or her own views the right to express those views freely and for the views of the child to be given due weight in accordance with the age and maturity of the child") and art. 14 (on freedom of thought, conscience, and religion). However, the Convention also ascribes to parents, extended family, and communities "rights" to direct and guide children's exercise of any rights they have (art. 5), which this article will show to be problematic. For a child-centered critique of the UNCRC, see Dwyer, "Inter-Country Adoption and the Special Rights Fallacy."

2 See Howard, "Deciding for the Incompetent" (analyzing more and less respectful ways of deciding on behalf of persons with some agential capacity that is nevertheless insufficient to be characterized as autonomous).

3 Peter Vallentyne writes: "An individual is psychologically autonomous just in case (roughly) she has a sufficiently good capacity for rational reflection and revision of her beliefs, desires, and intentions" ("Libertarian Perspectives on Paternalism," 182).

Anglo-American legal systems have long placed together these groups of nonautonomous persons and adopted a similarly solicitous disposition toward all of them, authorizing state actors to exert protective control and supervision over aspects of their lives that for autonomous persons would generally be outside the state's proper authority—that is, would be, by right, matters of self-determination.[4] These legal systems have done so under the banner of *parens patriae*, a term meaning parent of the country and connoting an inherent authority and duty of government to exercise protective care and control as to persons unable to guard or promote their own welfare. Yet the precise nature of the state's role *qua parens patriae* is obscure. This article aims to clarify how we should understand it when it is appropriately deployed—in particular, whether and how *parens patriae* differs from the more familiar "police power" role that states fulfill when they act as the agent for society as a whole, promoting public goods and resolving conflicts between rights. (The word 'police' in this standard characterization is misleading; 'police power' connotes not just law enforcement but any state action in service to society, including, for example, building highways.)

Which model of decision-making state actors adopt for *parens patriae* intervention has important real-world consequences for persons deemed nonautonomous. This article focuses on the specific question whether the best model allows interests of other individuals or of society to influence outcomes or whether instead, it requires an exclusive focus on the welfare of the nonautonomous persons. It aims to answer this first by undertaking conceptual analysis aimed at identifying discrete and coherent conceptions of the *parens patriae* role and then by offering normative reasons for choosing one among those conceptions.

Philosophers and legal scholars have paid little direct attention to this question. They have addressed it only in an indirect and narrow way by theorizing about parental entitlement to possess legal power over children's lives, without acknowledging and defending the implicit view of the state's role in children's lives that their theory entails—namely, that the state appropriately exerts power over certain aspects of children's lives in order to serve interests of other persons (parents).[5] This article shifts focus to the state and its position in the lives of

4 See *E. (Mrs.) v. Eve (Re Eve)* [1986] 2 SCR 388 (Canada), § 72: "In early England, the *parens patriae* jurisdiction was confined to mental incompetents, but its *rationale* is obviously applicable to children and, following the transfer of that jurisdiction to the Lord Chancellor in the seventeenth century, he extended it to children under wardship."

5 Philosophers generally agree that defense of parents' rights must rest on parental interests. Common arguments among nonphilosophers that parents have a moral right to control children's lives because that is good for children rest on lack of understanding of the nature of rights and failure to distinguish entitlement from authority. Rights are entitlements

children and other nonautonomous persons. And so, in part, it casts the parents' rights debate in a new light that should yield fresh insights. Rather than beginning with parents' perspectives and desires, and rather than rehearsing and critiquing the philosophical literature on parents' rights, this article focuses on and develops a theory of the state's relationship to children.[6] At the same time, it broadens the scope of theorizing about the state's treatment of nonautonomous persons to encompass also any adults in that category. This should foster more principled reasoning than one finds in writings limited to child rearing. From this higher vantage, the article aims to clarify *parens patriae*'s practical scope, how it ought to be conceptualized, and what aims it may properly serve.

Section 1 presents a taxonomy of relevant state actors and the types of decisions they make that impact nonautonomous persons. Section 2 briefly recounts the history of *parens patriae* in the Anglo-American legal tradition. Section 3 describes the diversity of conceptions of the *parens patriae* role that judges and legal scholars have deployed in the modern era. The next two sections form the core of the analysis. Section 4 assesses which conceptions of the *parens patriae* role are coherent and differentiates it from the state's more familiar police power role. Section 5 illustrates how choice among coherent conceptions on normative grounds depends on which moral outlook one believes underlies existing constraints on state exertion of power over the private lives of autonomous individuals. It ultimately recommends a conception in which the state *qua parens patriae* acts as fiduciary for nonautonomous individuals, obligated to serve only those individuals and to effectuate their rights (choice-protecting or interest-protecting as appropriate), as most consistent with the prevailing, deontological view of why the state respects individual rights. Not everything the state does that impacts nonautonomous persons is done *qua parens patriae*. But when the state does act in that role—that is, when it takes over decisions in an individual's life that for autonomous persons are a

correlative to duties owed to right holders, and duties are owed to persons whose interests are the moral foundation for their existence. See Sreenivasan, "Duties and Their Direction," 484; and Kramer, "Refining the Interest Theory of Rights," 32. An argument that begins with children's interests should therefore end with children's rights, not with rights of anyone else. Children might have a right that their parents possess substantial *authority* in their lives, but that is different from saying parents are *entitled* to possess such authority, such that the state owes a duty *to the parents* to confer that authority. Similarly, we say incompetent adults have a right that their guardians (as opposed to, for example, the state's department of social services) possess authority over certain aspects of their lives, but no guardian has a right to possess such authority. Likewise, criminal defendants have a right that their attorneys possess substantial authority in legal proceedings, but the attorneys have no entitlement to such authority.

6 For review and critique of parents' rights arguments, see Dwyer, "Deflating Parents' Rights."

matter of self-determining right (e.g., to undergo a medical procedure)—the state is, just like a guardian or other private fiduciary, constrained by the fiduciary duty of loyalty to serve only the welfare of that individual.

Finally, section 6 explores implications of adopting this fiduciary conception of *parens patriae*. Doing so is consistent with the law's current treatment of adults who lack legal competence, but it calls into question prevailing intuitions today about parents' normative position in decision-making about children's lives and the many legal rules that rest on those intuitions, and it poses a new kind of challenge to theorists who endeavor to defend parents' rights. There might be morally sound reasons for treating different groups of nonautonomous persons differently in the respect this article addresses—that is, as to whose interests count in state decision-making about certain aspects of their lives. However, this article's analysis generates conclusions as to the state's response to lack of autonomy in general, so it places on anyone who insists on different treatment of some group (e.g., children) the burden of presenting such reasons and of showing in a principled way why the general conclusions should not apply to that group.[7]

1. THE WHO AND WHAT OF *PARENS PATRIAE*

The state acts through many entities—executive officers, administrative agencies, legislatures, courts, and so on. Decisions by any of those entities can impact nonautonomous persons. Some decisions do not target nonautonomous persons specifically but rather concern the population generally—for example, decisions to improve highways or to prohibit theft. As to those state actions, nonautonomous persons stand in the same relation to the state as autonomous persons, and such actions are regarded as ordinary aspects of police power governance of a society, not part of the state's *parens patriae* authority. Other state decisions, though, do target persons deemed nonautonomous, aiming to disable, protect, or benefit them in light of their lesser capacities or greater vulnerability, in some way that the state ordinarily does not do with autonomous persons—for example, ordering that they receive certain medical treatment or education. State officials frequently invoke the state's *parens patriae* role as the basis for their authority to make this latter sort of decision for children or incompetent adults.

Some targeted decisions or acts establish uniform rules and prescribe the same treatment for large numbers of nonautonomous persons. These could be

7 See Munby LJ in *Re SA (Vulnerable Adult with Capacity: Marriage)* [2005] EWHC 2942 (Fam) (Munby J) § 37: "It is now clear . . . that the court exercises what is . . . a jurisdiction in relation to incompetent adults which is for all practical purposes indistinguishable from its well-established *parens patriae* or wardship jurisdictions in relation to children."

executive orders, legislation, or administrative regulations. Examples include a health department directive that all residents of Alzheimer's units in care facilities receive periodic medical exams and legislation prescribing which subjects children should learn in schools. Other state decisions are specific to individuals—for example, a court order of surgery for a mentally disabled adult or a child protection agency's decision to remove a child from the custody of neglectful parents. Legislative prescriptions do not inherently differ in content from individualized decisions, however; the state can make the same kind of decision by either means. For example, states place infants into legal parent-child relationships both by legislation (e.g., via presumptions of maternity based on giving birth and presumptions of paternity based on being married to the mother) and by individualized court decisions (e.g., via adoptions and challenges to the marital presumption of paternity). Allocation of decisions to different types of state actors instead typically reflects efficiency judgments, depending on how common and easily established are the determinative facts.

Both state decisions that address a class of persons and those focused on a particular individual can either directly dictate some aspect of nonautonomous persons' lives, as in the examples above, or delegate direct decision-making to a private entity. As to the latter, legal systems have statutes conferring broad presumptive powers on guardians and conservators for incompetent adults and on parents of minor children.[8] But in addition, courts sometimes tailor the powers of such caretakers to individual circumstances (e.g., authorizing a conservator to sell the ward's major assets or dividing authority over a child's education in a particular way between divorced parents).[9]

Finally, state decisions and actions that target nonautonomous persons can pertain to a vast range of substantive matters: formation and dissolution of legal relationships; choices as to residential location, medical care, education, and finances; avoidance of disturbing speech or conduct by custodians; and so on. Importantly, all possible decision contexts have analogues in the lives of autonomous persons—that is, situations in which autonomous persons make similar decisions for themselves. This is more readily apparent with some practical matters. For example, my making decisions about my own health care is a clear analogue to the state's actions of mandating some forms of health care

8 See, e.g., Protection of Personal and Property Rights Act 1988 (New Zealand) (PPPRA), § 18(2) (welfare guardian for incompetent adult); Uniform Probate Code (U.S.) § 72-5-427 (UPC) (conservatorship); and Arizona Revised Statutes § 1-601 (declaring broadly the "liberty of parents to direct the upbringing, education, health care and mental health of their children").

9 See, e.g., Massachusetts Gen. Laws, UPC § 72-5-425 (tailoring of conservator powers); and Florida Statutes § 61.13 (court discretion regarding legal custody).

for all children, such as immunizations and medical treatment in the event of injury. Less familiar perhaps are analogues to the state's creation of legal caretaking relationships for nonautonomous persons (i.e., guardian-ward and parent-child) and delegation of decision-making power to the caretakers. But autonomous adults do make similar decisions. They can choose for themselves who will serve as their guardian or make medical decisions for them in the event of their incapacity (in an advance directive or health care proxy).[10] They can select someone to manage their finances should they lose competence or even while they remain competent (e.g., while they travel, undergo major surgery, or devote their attention to other things) by creating a trust with someone else as trustee, nominating someone for conservatorship, or conferring a power of attorney.[11] The caretaker or agent whom one selects can, like a parent, be someone with whom one shares a close family relationship and daily life, such as a spouse, sibling, or adult offspring. Indeed, it can be one's parent.

The point of establishing that state decisions specifically concerning nonautonomous persons all have analogues in decisions autonomous persons make for themselves is to motivate consideration below of whether the former, which state actors typically characterize as carried out in their *parens patriae* capacity, can and should be guided by principles underlying the law's treatment of autonomous adults' decisions about their own lives in those same realms. Before delving into that question, the next two sections provide historical background and contemporary context.

2. ORIGINS

Scholarly writing on the state's *parens patriae* authority over nonautonomous persons is predominantly historical.[12] The origins are somewhat unclear.[13]

10 See, e.g., PPPRA (New Zealand), § 12(7); Patients Property Act (PPA), Rev. Stat. British Columbia [1996] (PPA), Ch. 349, § 9; Massachusetts Gen. Laws, UPC § 72-5-305(b) ("Unless lack of qualification or other good cause dictates to the contrary, the court shall appoint a guardian in accordance with the incapacitated person's most recent nomination in a durable power of attorney"); and Conn. Gen. Stat. § 19a-575a(a) ("Any person eighteen years of age or older may execute a document that contains health care instructions, the appointment of a health care representative, the designation of a conservator of the person for future incapacity").

11 See, e.g., PPA (British Columbia), Ch. 349, § 9; UPC § 72-5-410.

12 See, e.g., Seymour, "*Parens Patriae* and Wardship Powers"; Payton, "The Concept of the Person in the *Parens Patriae* Jurisdiction over Previously Competent Persons"; and Custer, "The Origins of the Doctrine of *Parens Patriae*."

13 *Re Eve* [1986] 2 SCR 388 (Canada) § 32: "The origins of the Crown's *parens patriae* jurisdiction over the mentally incompetent . . . is lost in the mists of antiquity."

The most plausible view seems to be that it originated in seventeenth-century England with transfer of jurisdiction over "infants," "lunatics," and "imbeciles" from the Court of Wards and Liveries (which was concerned principally to protect the Crown's interest in feudal tenancies when nonautonomous persons inherited land) to the Chancery Court, "Keeper of the King's Conscience" (which was charged generally with effecting justice for the king's subjects, with greater flexibility and discretion than the common law courts).[14] Though initially concerned principally with appointing guardians and preserving ward's property, the *parens patriae* function came to encompass also protective intervention to stop abuse and neglect by custodians, whether parents or appointed guardians.[15] The concept of neglect expanded over time to include failure to secure not just necessities of bare survival but also developmental goods such as education and preventive health care.[16] Still later, it extended to custody disputes between parents.[17]

14 Custer, "The Origins of the Doctrine of *Parens Patriae*," 206; and Seymour, "*Parens Patriae* and Wardship Powers," 175–76. From at least the thirteenth century, the English monarch assumed "wardship" of incompetent persons, but primarily or solely those who owned property, in order to extract rents. Australian jurist Paul Brereton observes that "the feudal wardship system was a rudimentary predecessor" to the altruistic *parens patriae* jurisdiction ("The Origins and Evolution of the *Parens Patriae* Jurisdiction," 4). Cf. *In re Spence* (1847) 2 Ph 247 (England), Op. Lord Cottenham: "The cases in which this court interferes on behalf of infants are not confined to those in which there is property. . . . This court interferes for the protection of infants, *qua* infants, by virtue of the prerogative which belongs to the Crown as *parens patriae*."

15 Joseph, "Inherent Jurisdiction and Inherent Powers in New Zealand," 227: "Over time, wardship became procedurally and substantively connected with the *parens patriae* jurisdiction. It lost its connection with property and became purely protective in nature."

16 *Re Eve* [1986] 2 SCR 388 §§ 35ff (identifying cases in Canada and England invoking *parens patriae* to support court-ordered medical procedures) and § 74 ("It can be invoked in such matters as custody, protection of property, health problems, religious upbringing and protection against harmful associations"); *Re Jules* [2008] NSWSC 1193 (New South Wales) 14 ("It has been said that no limit has ever been set to it and that it extends as far as is necessary for the protection and education of children"); *Re X (a Minor)* [1975] 1 All ER 697, 703 (England) ("The court has power to protect the ward from any interference with his or her welfare, direct or indirect"); and Thomas, "Limitations on *Parens Patriae*," 57 ("Ever since the mid-19th century, state legislatures have relied on the *parens patriae* doctrine in enacting compulsory-education laws").

17 *Henderson v. Henderson*, 91 A.2d 747, 749–50 (NJ 1952) ("The power of the former Court of Chancery, to which the Superior Court succeeded as *parens patriae*, is firmly established in our jurisprudence. . . . In dealing with the custody and control of infants the touchstone of our jurisprudence is their welfare and happiness"); and Bishop, *Commentaries on the Law of Marriage and Divorce and Evidence in Matrimonial Suits*, § 636.

All branches of government in England and in countries that inherited its legal system have participated in fulfilling the *parens patriae* role.[18] Adoption of the function originally was typically not by explicit authorization in government charters such as constitutions, nor with public deliberation. Rather, courts, legislative bodies, and executive officers simply continued a familiar, taken-for-granted state practice based upon a protective authority viewed as inherent to enlightened government.[19] Over the centuries, though, many jurisdictions have codified the authority.[20] Other legal traditions in the Western

18 See Graham, "*Parens Patriae*," describing its application in Australia, Canada, and New Zealand. See also Payton, "The Concept of the Person in the *Parens Patriae* Jurisdiction over Previously Competent Persons," 618 (the *parens patriae* function "was incorporated into the equity jurisdiction of the newly independent American states with no recorded objection"); *In re DS*, 763 N.E.2d 251, 261 (Ill. 2001) ("Each branch of government has concurrent powers and responsibilities that are in the nature of *parens patriae*"); and *In re J. J. Z.*, 630 A.2d 186, 193 (DC 1993) ("The statute also places within the power of the executive branch, through the Corporation Counsel, *parens patriae* responsibilities to the child").

19 See Seymour, "*Parens Patriae* and Wardship Powers," 162: "The first [US] case to invoke the *parens patriae* power in the public custody context was *Ex parte Crouse*, 4 Whart. 9 (PA 1839). The court cited no authority, and simply asked 'may not the natural parents, when unequal to the task of education, or unworthy of it, be superseded by the *parens patriae*, or common guardian of the community?'" See also *Re Eve* [1986] 2 SCR 388 (Canada), § 65 ("That jurisdiction is based on the inherent equitable power of the courts to act in the best interests of the mentally incompetent person"); *Late Corporation of Church of Jesus Christ of Latter-Day Saints v. United States*, 136 US 1, 57–58 (1890) ("This prerogative of *parens patriae* is inherent in the supreme power of every state, whether that power is lodged in a royal person, or in the legislature . . . for the prevention of injury to those who cannot protect themselves. . . . [T]his beneficent function has not ceased to exist under the change of government from a monarchy to a republic, but it now resides in the legislative department, ready to be called into exercise whenever required for the purposes of justice and right"); and *Vidal v. Girard's Ex'rs*, 43 US 127, 144, 167–68, 195 (1844) ("From the time of Augustine, the common law had been undergoing changes to suit the spirit of the age, but the revealed law was a part of it all the time. . . . To this same great source we owe the idea of a paternal power in the state—a *parens patriae*—not the king, nor the chancellor, but a power existing somewhere to take care of the sick, the widow, and the orphan").

20 See, e.g., the Mental Capacity Act (MCA) 2005 (UK) ("An act done, or decision made, under this Act for or on behalf of a person who lacks capacity must be done, or made, in his best interests"); Scottish Adults with Incapacity Act (2000); Chancery Act, Rev. Stat. 1951 (Prince Edward Island), c. 21, sec. 3; and Judicature Act 1908 (New Zealand), sec. 17. See also Donnelly, "Inherent Jurisdiction and Inherent Powers of Irish Courts," 134: "Section 47 of the Regulation of Commission in Lunacy Act, 1853 established a Court of Protection entrusted with functions to manage the affairs of legally incapacitated adults. In Ireland, the state's wardship jurisdiction with respect to persons of unsound mind was granted a separate statutory basis by virtue of the Lunacy Regulation (Ireland) Act, 1871." See also Kindred, "God Bless the Child," 526: "The *parens patriae* power has been recognized from earliest times in the United States as well and now is largely governed by state statutes."

world assign a similar role to the state.[21]

This background is just to establish long and widespread acceptance that an enlightened society will assign to the state a function of protecting persons unable to safeguard their own interests.[22] Further, that in doing so the state exercises a power or jurisdiction that is special, different in some way from state action in contexts involving only autonomous persons, including in its scope, insofar as it entails intervening in areas of their life ordinarily considered private and not subject to state control.[23] How exactly it is different in its nature and operation, though, has never been a subject of close study. This article's aim is not to establish a prevailing understanding of the role among state actors historically but rather to establish normatively what view they should adopt.

3. ALTERNATIVE MODERN CONCEPTIONS

Current conceptualization of the state's *parens patriae* role as protector of nonautonomous persons is, in Anglo-American jurisprudence, inconsistent in material respects across instances of its application by various legal actors.[24] In judicial decisions and legal scholarship, one finds two seemingly disparate understandings of that role.

First, many ostensibly treat *parens patriae* as a special fiduciary role in which the state serves as agent not for society as a whole, as it does in its police power role.[25] Rather, it serves as agent solely for nonautonomous individuals, as a private fiduciary would.[26] A distinctive feature of fiduciary roles is a duty of

21 See Merkel-Holguin et al., *National Systems of Child Protection* (collected papers describing child protection systems in Australia, Canada, England, France, Germany, Ireland, Israel, the Netherlands, South Korea, and Switzerland); and Holland, *The Law Relating to the Child, Its Protection, Education, and Employment* (describing child protective laws in Spain, Germany, France, and Italy).

22 *Late Corporation*, 136 US at 58: "Take this away and we become a nation of savages." See also Legarre, "The Historical Background of the Police Power," 764: "The doctrine of *parens patriae* is present in some form and under some name in every reasonable legal system, for no reasonable state would wish to violate the basic ethical imperative of looking after children deserted by their parents."

23 Payton, "The Concept of the Person in the *Parens Patriae* Jurisdiction over Previously Competent Persons," 641: "The *parens patriae* jurisdiction . . . is fundamentally unlike other powers of the state."

24 See *In re Gault*, 387 US 1, 16 (1967), saying of *parens patriae* that "its meaning is murky."

25 Gold and Miller, "Introduction," 1: "There is even a storied tradition of thinking of the authority of the state in fiduciary terms."

26 A large literature has developed in recent years debating whether the fiduciary model of agency in private law can sensibly be exported to government in its police power capacity, serving society as a whole, given the great diversity and conflict of interests across

loyalty, owed to the person for whom the agent is a fiduciary, entailing exclusive devotion to that person's welfare.[27] The Supreme Court of Canada advanced this fiduciary conception with unusual explicitness in its seminal *Re Eve* case in 1986, rejecting "an application by a mother for permission to consent to the sterilization of her mentally retarded daughter who also suffered from a condition that makes it extremely difficult for her to communicate with others":

> The *parens patriae* jurisdiction is … founded on necessity, namely the need to act for the protection of those who cannot care for themselves.… The courts have frequently stated that it is to be exercised in the "best interest" of the protected person, or again, for his or her "benefit" or "welfare." It must be exercised in accordance with its underlying principle. Simply put, the discretion is to do what is necessary for the protection of the person for whose benefit it is exercised.… The discretion is to be exercised for the benefit of that person, not for that of others.… This is particularly so in cases where a court might be tempted to act because failure to do so would risk imposing an obviously heavy burden on some other individual.… One may sympathize with [Eve's mother].… But the *parens patriae* jurisdiction cannot be used for her benefit. Its exercise is confined to doing what is necessary for the benefit and protection of persons under disability like Eve.… So we are left to consider whether

groups and individuals within a society. See Leib et al., "Mapping Public Fiduciary Relationships"; and Criddle et al., eds, *Fiduciary Government*. That problem does not arise when the state acts solely as agent for an individual, as courts often do—for example, in deciding whether to order a medical procedure for an incompetent adult. "In some cases public officials undertake fiduciary duties, as where the state exercises custodial power over children.… But exercising power of this sort relative to an ascertainable beneficiary is very different form exercising power for the sake of the public" (Criddle et al., "Introduction," 17). Somewhat different is legislation as to some aspect of life for a large group of nonautonomous persons, if there can be a conflict of interest among them (e.g., if two or more co-reside or are in the same classroom). But there is also a private law analogue to that situation, in trust law, where the fiduciary duty of loyalty cashes out in a duty of impartiality across beneficiaries coupled with the universal proscription of self-dealing or aiming to serve nonbeneficiaries. See the US Uniform Law Commission's Uniform Trust Code § 803. And much legislation concerning nonautonomous persons does not present such intergroup conflicts (e.g., a patients' bill of rights or teacher qualification requirements for schools).

27 Gold and Miller, "Introduction," 5: "The duty of loyalty is one of the most prominent features of fiduciary law." See also Bogert et al., *The Law of Trusts and Trustees*, § 543: "Perhaps the most fundamental duty of a trustee is the trustee's duty of loyalty to the beneficiaries, often stated as the duty to act solely in the interests of the beneficiaries. This duty is sometimes stated as the rule of undivided loyalty. The trustee must administer the trust with complete loyalty to the interests of the beneficiary, without consideration of the personal interests of the trustee or the interests of third persons."

> the purposes underlying the operation are necessarily for Eve's benefit and protection. . . . Many [such persons who conceive a child], it is true, may have difficulty in coping, particularly with the financial burdens involved. But this issue does not relate to the benefit of the incompetent; it is a social problem. . . . Above all it is not an issue that comes within the limited powers of the courts, under the *parens patriae* jurisdiction, to do what is necessary for the benefit of persons who are unable to care for themselves. . . . Accordingly, the procedure should never be authorized for nontherapeutic purposes under the *parens patriae* jurisdiction. . . . The fact that others may suffer inconvenience or hardship from failure to do so cannot be taken into account. The Crown's *parens patriae* jurisdiction exists for the benefit of those who cannot help themselves, not to relieve those who may have the burden of caring for them.[28]

Likewise, in a case involving a minor and a petition to order lifesaving medical treatment, the Supreme Court of New South Wales stated:

> In a case such as the present, the Court is not balancing the interests of the individual against broader public or governmental interests. . . . All humans affect others and are affected by a myriad of relationships. . . . However, the *parens patriae* jurisdiction is not used for or directed to the benefit of parents or others related to or connected with a child that is the subject of the Court's consideration. Its exercise is directed to, and in that sense circumscribed to, doing what is necessary for the benefit and protection of such child.[29]

Thus, if the state's *parens patriae* role is in the nature of a fiduciary for a nonautonomous individual, then presumptively state actors must, when making decisions for the individual in that capacity, adopt a singular aim of serving the dependent person's interests, to the exclusion of other considerations. The duty of loyalty would proscribe both self-dealing (serving state interests) and

28 *Re Eve* [1986] 2 SCR 388 (Canada) §§ 1, 73, 77, 82, 84, 86, 92.

29 *H v. AC* [2024] NSWSC 40 (New South Wales), §§ 54–55. See also *LS v. British Columbia* (2018) BCSC 255, § 30: "Courts have frequently stated that it is to be exercised in the 'best interest' of the protected person, or again, for his or her 'benefit' or 'welfare'. . . . This jurisdiction is to be exercised to protect children and other vulnerable individuals, not their parents." See also *JP v. British Columbia* (2015) B.C.S.C. 1216, concluding that social workers in a pediatric psychiatric facility "should not have to weigh what is best for the child on the scale with what would make the family happiest." See also *Rogers v. Okin*, 634 F.2d 650, 654, 657, 661 (1st Cir. 1980): "Following a determination of incompetency, state actions based on *parens patriae* interests must be taken with the aim of making treatment decisions as the individual himself would were he competent to do so."

aiming to serve interests of third parties (family members, etc.) when acting as agent for the dependent individual.[30]

Such surrogate decisions must take place, of course, just like autonomous persons' self-determining choices regarding similar matters, within the context of an existing distribution of societal resources determined by the market and by decisions the state renders in its police power capacity. The fiduciary conception of *parens patriae* does not entail that the state must commit all available state resources to the welfare of nonautonomous persons with no heed for the impact on other members of society. But once the state, wearing its police power hat, has determined a fair distribution of public resources across society, it must, when wearing its *parens patriae* hat, choose from among options then available to a nonautonomous person solely based on which will best serve that person's well-being. In other words, it steps into that person's shoes and acts on his or her behalf, with a presumption of only self-regarding motivations.[31] It would thus be improper to allow collective societal aims or preferences of other private individuals or groups directly to influence a state actor's decision-making about, for example, whether a nonautonomous person receives medical treatment that is available to them. That state actor would operate the same way we would expect of a private individual holding a medical power of attorney for an adult who has lost capacity to make decisions—that is, as a fiduciary under a duty of loyalty to the ward, with exclusive focus on the ward's interests, not balancing those interests against the desires of others or state interests.[32]

Outcomes in high-profile legal disputes concerning severely brain-damaged adults on life support have reflected this view, as have many cases involving petitions for sterilization (such as *Re Eve*), abortion, or other medical intervention.[33] Likewise with respect to children, many scholars and courts through-

30 See Uniform Trust Code § 802.

31 Miller, "Fiduciary Representation." Proxy altruism is discussed in detail in note 49 below.

32 Payton, "The Concept of the Person in the *Parens Patriae* Jurisdiction over Previously Competent Persons," 617: "Under the law that has governed the *parens patriae* jurisdiction ever since it was created . . . , a person whose powers of self-management have been taken from him by the state has a right that those who exercise the power to manage his affairs on his behalf do so in a fiduciary capacity. . . . The state takes jurisdiction only as a trustee: the jurisdiction has been designed to avoid vesting in the state any authority or incentive to act in a self-interested manner vis-à-vis the incompetent." The jurisdiction has been "designed from the beginning to be wholly fiduciary. . . . [T]he state's role has been exclusively that of trustee" (616).

33 See *K v. Minister for Youth and Community Services* [1982] 1 NSWLR 311 (New South Wales) (abortion); *In re Guardianship of Schiavo*, 851 So. 2d 182, 186–87 (Fla. Dist. Ct. App. 2003) (life support) ("this case is not about the aspirations that loving parents have for their children. It is about Theresa Schiavo's right to make her own decision, independent of her parents

out the English-speaking world have characterized *parens patriae* as calling for exclusive focus on the welfare of the child when state agencies must get involved in intimate aspects of their lives.[34] Others suggest a similar view by stating that a child's welfare is "paramount" or the like and not referring to any other interests as relevant.[35] Or they describe the state's role as that of a "wise

---

and independent of her husband. In circumstances such as these, . . . trial judges . . . serve as surrogates or proxies to make decisions about life-prolonging procedures. . . . [T]he trial judge must make a decision that the clear and convincing evidence shows the ward would have made for herself"); and *Cruzan by Cruzan v. Director, Missouri Department of Health*, 497 US 261, 286 (1990) (rejecting parents' assertion of a right to decide about life-sustaining care of their adult offspring, stating "we do not think the due process clause requires the state to repose judgment on these matters with anyone but the patient herself"); *In re Quinlan*, 355 A.2d 647, 661–62 (NJ 1976) ("We do not recognize an independent parental right of religious freedom to support the relief requested"). See also *O'Connor v. Donaldson*, 422 U.S. 563, 573–75 (1975) (stating that mere preferences of the public not to see mentally ill persons was not a constitutionally permissible consideration in commitment decisions); and *In re Terwilliger*, 450 A.2d 1376, 1382 (PA Super. Ct. 1982) ("In making the decision whether to authorize sterilization, a court should consider only the best interest of the incompetent person, not the interests or convenience of the individual's parents, the guardian or society").

34 See, e.g., Pomeroy, *A Treatise on Equity Jurisprudence*, § 1307, 330n1 ("the exercise of the jurisdiction depends on the sound and enlightened discretion of the court and *has for its sole object* the highest well-being of the infant"); *Smith v. Smith*, 26 Eng. Rep. 977, 977 (Ch. 1745) (England) ("It is not a profitable jurisdiction of the crown, but for the benefit of infants themselves"); *Wentzel v. Montgomery Gen. Hosp.*, 447 A.2d 1244, 1253 (Md. 1982) ("It is a fundamental common law concept that the jurisdiction of courts of equity over such persons is plenary so as to afford whatever relief may be necessary to protect the individual's best interests"); Harvard Law Review Association, "Developments in the Law," 1199 ("when the state acts as *parens patriae*, it should advance only the best interests of the incompetent individual and not attempt to further other objectives, deriving from its police power, that may conflict with the individual's welfare"), 1200 ("it should exercise the *parens patriae* power solely to further the best interests of the child"); and Hatcher, "Purpose vs Power," 171–72 ("State child welfare agencies exist to protect the interests, and the rights, of abused and neglected children. . . . The agencies serve in the nature of a fiduciary for children's rights").

35 *H v. AC* [2024] NSWSC 40 (New South Wales), § 47: "The inherent, protective jurisdiction of the Court exists for the protection of those (including, but not limited to, minors) who are unable to protect themselves. It accords paramountcy to the welfare of the person in need of protection." See also *J. v. C.* [1970] AC 668, 697 (UK) (Guest LJ) ("the law administered by the Chancery Court as representing the Queen as *parens patriae* never required that the father's wishes should prevail over the welfare of the infant. The dominant consideration has always been the welfare of the infant"); *Re Frances and Benny* [2005] NSWSC 1207, 17 (Young CJ in Eq) ("In exercising that [*parens patriae*] jurisdiction the court's concern is predominantly for the welfare of the person involved"); *In re J. J. Z.*, 630 A.2d 186, 193 (DC 1993) ("this court has recognized the longstanding principle that 'in a civil proceeding predicated on alleged child neglect or abuse, the best interest of the child is the paramount consideration.... Neglect statutes authorizing state intervention on a child's

parent" or guardian.[36]

However, there is a competing conception with some currency, albeit one not so clearly articulated, mostly implicit. Other sources appear to treat *parens patriae* decision-making as a matter of simply paying special heed in some way to the interests of dependent persons while also letting preferences or interests of other individuals or of society as a whole influence state decisions about a nonautonomous person's life. On this *special heed conception*, it seems state actors aim simply to ensure nonautonomous persons' interests receive direct attention and carry some weight—perhaps heightened, perhaps equal (this is never clear)—along with other individuals' interests and even collective societal aims that might be impacted by decisions, in some sort of balancing if those several interests do not all align.[37] The state does not act as agent exclusively

behalf … should be liberally construed to enable the court to carry out its obligation as *parens patriae*'"); *In re Karwath*, 199 N.W.2d 147, 150 (Iowa 1972) ("Our paramount concern for the best interests and welfare of the children overrides the father's contention that absolute medical certitude of necessity and success should precede surgery"); *Henderson v. Henderson*, 91 A.2d 747, 750n18 (NJ 1952) ("the touchstone of our jurisprudence is their welfare and happiness"); *Vannucchi v. Vannucchi*, 272 A.2d 560, 563 (NJ App. Div. 1971) ("In the exercise of their *parens patriae* jurisdiction, our courts look always to the protection of the child's best interests—the happiness and welfare of a child are paramount in determining custody").

36 E.g., *Re Jules* [2008] NSWSC 1193, 16 ("In exercising this jurisdiction, the Court endeavours to act as would a wise parent.… The Court may in place of the parents make those decisions which it considers appropriate in the best interests of the child"); *Re J* [1992] 4 All Eng. L. Rep. 614 (characterizing the court in medical treatment cases as "adopting the standpoint of reasonable parents who had the child's best interests at heart without regard to their own interests"); *Rogers v. Okin*, 634 F.2d 650, 654, 657, 661 (1st Cir. 1980) (US) ("The concept of *parens patriae* … developed with reference to the power of the sovereign to act as 'the general guardian of all infants, idiots, and lunatics'"); *Kicherer v. Kicherer*, 400 A.2d 1097, 1100 (MD 1979) ("In reality the court is the guardian; an individual who is given that title is merely an agent or arm of that tribunal in carrying out its sacred responsibility"); *R. v. Gyngall* [1893] 2 QB 232, 239 ("It was a paternal jurisdiction … as being the guardian of all infants, in the place of a parent, and as if it were the parent of the child"); and Kindred, "God Bless the Child," 526 ("This state power, known as the *parens patriae* doctrine, in essence, gives the state authority to serve as a substitute parent and ultimate protector of children's interests"). Other courts have stated that the state's *parens patriae* power can extend beyond that accorded parents in a given jurisdiction to include such things as authorizing civil commitment, certain medical treatments, and underage marriage. See, e.g, *Re R.* [1991] 4 All ER 177, 186 (England) (involuntary administration of antipsychotic medication).

37 See, e.g, *Troxel v. Granville*, 530 U.S. 57, 88 (2000) (Stevens, J., dissenting) (asserting, in a case concerning parental control over children's visitation with grandparents, that "a parent's interests in a child must be balanced against the State's long-recognized interests as *parens patriae* … and, critically, the child's own complementary interest in preserving relationships that serve her welfare and protection"); *Matter of Weberlist*, 360 N.Y.S.2d

for the individual; this is what distinguishes this special heed conception from the fiduciary conception. For example, if the issue is with whom a person will live, whereas a competent adult can choose not to live with someone whose religiously grounded speech upsets them, giving no consideration to that other person's desire to cohabit or right to free speech, on the special heed conception of *parens patriae*, the state making the same sort of decision about cohabitation for a mentally disabled adult would permissibly take into account the disappointment a potential guardian would feel if denied the role and the impact it could have on that person's sense of religious or expressive freedom.[38] A competent adult can choose to undergo cosmetic surgery without regard for the feelings or moral beliefs of parents, other family members, or the public, but on the special heed view, it is appropriate for the state to aim to satisfy such third parties when legislating as to whether minors or adults under a guardianship can undergo any such procedures; the state must simply also attend in some way to the minors' or wards' own interests.

This special heed view thus appears implicitly to suppose that, given nonautonomous persons' inability to advocate for themselves, absent assertion of *parens patriae* authority, their interests would receive no consideration at all in legal decisions about their lives or in private actors' decisions that impact them. The state might cite its *parens patriae* role to explain why it has gotten involved in some aspect of a nonautonomous person's life when no one asked it to do

783, 786 (1974) ("the state must intervene in order to protect an individual who is not able to make decisions in his own best interests. The decision to exercise the power of *parens patriae* must reflect the welfare of society, as a whole, but mainly it must balance the individual's right to be free from interference against the individual's need to be treated, if treatment would in fact be in his best interest"); *Ex parte Wallace*, 190 P. 1020, 1022 (NM 1920) ("in all cases the state is *parens patriae* to the child, and it has power to, by legislation, control the right of the child to inherit, to take it from its parents, and give it into custody of others, to determine what is for the best interests of the child, and that which will promote the welfare of the state"). See also Seymour, "*Parens Patriae* and Wardship Powers," 186 ("The fact that the need to treat the welfare of the child as the paramount consideration is a distinguishing feature of some types of proceedings does not mean that these proceedings can be conducted without regard to the interests of others who might be affected"); and Weithorn and Reiss, "Providing Adolescents with Independent and Confidential Access to Childhood Vaccines," 797 (characterizing the effect of *parens patriae* action as "triggering a higher level of protection from the state" for nonautonomous persons).

38 *Shepp v. Shepp*, 588 PA 691 (2006), holding that a father had a constitutional right against being denied custody of his daughter because of his efforts to impress Mormon views on her, including the view that she should enter into a plural marriage and choose a life of service to a husband (even though that expression of views had caused his wife to petition for divorce), unless his expression was proven to "jeopardize the physical or mental health or safety of the child, or have a potential for significant social burdens" (706). See also Volokh, "Parent-Child Speech and Child Custody Speech Restrictions."

so—for example, when it intervenes to block choices or conduct by guardians or parents that it deems detrimental.

The basic distinction between the two conceptions, then, is that the fiduciary conception does not allow interests other than those of the ward to influence decisions directly, whereas the special heed conception does, even when those other interests are antithetical to those of the ward. Invocation of one conception or the other is typically undefended, and invocation of the special heed conception typically unreflective. Yet in numerous contexts, state decisions about central aspects of nonautonomous persons' lives are profoundly important, and the model of *parens patriae* adopted could determine outcomes. This is clearest when a legislature or court says explicitly that a best interests finding is not enough to override competing interests.[39] Among the most important state decisions for children is choice of legal parents. Though generally unrecognized, state laws that confer initial legal parenthood on birth parents, almost without exception (one exception being anonymous birth, another being denial of paternity to rapists), rather than attempting to screen out those manifestly unfit to care for a child, reflect a legislative choice. Even if one believes the state morally must make that choice, the fact is that the state is making a choice, and that choice greatly influences the life courses of children. State decisions about the residence and care of mentally disabled or mentally ill adults, when they cannot live with family, are similarly impactful. Less momentous decisions can have a cumulative effect in shaping lives and determining experiences.

39 A salient example in the US is the Indian Child Welfare Act, which explicitly aims to serve tribes' interests in sustaining membership by channeling to tribal lands children in mainstream society whom the state must place in foster care or adoption because of parental maltreatment. Most Supreme Court justices acknowledged in a recent decision that the act frequently operates to the children's detriment—in particular, when placement preferences force termination of a child's attachment relationship with long-term non-Indian caretakers. *Haaland v. Brackeen*, 143 S. Ct. 1609, 1622 (2023) (majority). See also JJ Thomas and Alito dissenting (1662). See also the Code of Federal Regulations, 25 CFR 23.143(c) (stating the position of the Bureau of Indian Affairs that a best-interests determination does not constitute "good cause" to deviate from adoption preferences). See also *Palmore v. Sidoti*, 466 U.S. 429 (1984) (subordinating child welfare concerns in a custody dispute between parents to the state aim of eliminating racism); and *People in Int. of D.L.E.*, 614 P.2d 873, 875 (CO 1980) (refusing to order treatment for seizures, out of deference to mother's religious beliefs, unless and until the child was in imminent danger of death; a best interests finding was insufficient).

4. CONTRASTING *PARENS PATRIAE* WITH OTHER STATE FUNCTIONS

Determining which conception is normatively preferable might begin by situating *parens patriae* within the full range of state functions. Though most people might think of the state as always operating in the same way, in a police power capacity, the state in any society inevitably operates in many capacities. As noted, when wearing its police power hat, the state acts within its borders as agent for all members of society, promoting the collective good and facilitating social interaction. It legislates and in other ways acts based upon consideration of everyone's interests, giving equal weight to like interests of all persons, balancing interests that conflict. It promotes general welfare and guards against harms to property and person.[40] Paradigmatic exercises of police power include managing the economy, building infrastructure, creating a social safety net, enacting and enforcing statutes criminalizing violence and theft, protecting public goods like the environment, creating rules for orderly movement in public and for transacting, resolving property or contract disputes, and promoting public health (e.g., mandating masks, funding medical research).[41]

However, the state regularly acts in other capacities. At one level or another, government can be an actor in international affairs, a competitive business operator, an employer, a jailer, a speaker in commercial advertising or the public square, a service provider, a funder of private service provision, a party to contracts, or an educator. Each of these roles has its own scope of authority and concerns, powers and duties, and governing norms and constraints. In the United States, different lines of constitutional doctrine pertain to those various functions.[42] For example, individual rights that constrain government when it acts in its police power role might do so less or not at all when it operates in one

40 Brady, "Turning Neighbors into Nuisances," 1659: "[police power] demarcates the boundaries within which the government can affirmatively regulate for the health, safety, and welfare of its citizens, limiting some rights to protect the greater good."

41 Legarre, "The Historical Background of the Police Power," 774: "'police power' refers to the authority of the states for the promotion of public health, public safety, public morals, and public welfare."

42 *Enquist v. Oregon Department of Agriculture*, 553 U.S. 591, 598 (2008) ("there is a crucial difference, with respect to constitutional analysis, between the government exercising 'the power to regulate or license, as lawmaker,' and the government acting 'as proprietor, to manage [its] internal operation'"); *Legal Services Corporation v. Velazquez*, 531 U.S. 533, 541 (2001) ("viewpoint-based funding decisions can be sustained in instances in which the government is itself the speaker"); *Hazelwood School District v. Kuhlmeier*, 484 U.S. 260, 266 (1988) ("A school need not tolerate student speech that is inconsistent with its 'basic educational mission,' even though the government could not censor similar speech outside the school"); *Lehman v. Shaker Heights*, 418 U.S. 298 (1974) (public bus is not a public forum for free speech purposes because it is a commercial enterprise); and *Fulton*

of the other roles.[43] This is especially so when a state function approximates that of a private actor; for example, government employees have diminished free speech rights in relation to the government *qua* employer.[44] Some might find it hard to grasp this, reacting by saying "But it's the state!" and struggling to understand how any part of the Constitution could ever not apply to state actors. It is because the Constitution was designed for the core police power function, not for everything the state might do. The *parens patriae* role might be another in which government is not bound by the same norms, because it is acting in a quasi-private capacity. At the same time, government might be under additional *duties* to individuals in any of these non–police power roles—for example, duties *qua* employer that it owes to employees, such as to pay a salary, create healthy working conditions, and provide health insurance.

How, then, to reason to a conclusion about the particular nature of the *parens patriae* role within the range of state functions? We might begin by considering whether and how it differs conceptually from the core police power function, with which it is most often conflated or conjoined. This function has always been the primary focus of political theorists, who typically presuppose the people on whom governments act are autonomous.

Relative to the police power, one possible view of the *parens patriae* role is that it is not distinct; it is simply what we call exercise of police power when nonautonomous persons are centrally involved.[45] This would be consistent with the special heed conception. Its invocation might simply serve as a corrective to a tendency to do police power decision-making badly when nonautonomous persons are impacted, leaving them out of a cost-benefit equation that is supposed to include everyone's interests and objectively weigh each. Invoking the state's *parens patriae* responsibility could be simply a reminder: "Don't forget that intellectually disabled persons' lives matter too!"

---

v. City of Philadelphia, 320 F.Supp.3d 661 (ED PA 2018) (conditions impinging religious freedom in "a state contract for … services" are not subject to constitutional challenge).

43 See Kalb, "Gideon Incarcerated," 111, discussing greater judicial deference to governments in their operation of prisons.

44 *Garcetti v. Ceballos*, 547 U.S. 410, 421 (2006): "when public employees make statements pursuant to their official duties, the employees are not speaking as citizens for First Amendment purposes, and the Constitution does not insulate their communications from employer discipline."

45 See *In re Spence* (1847) 2 Ph. 247, 252; 41 E.R. 937, 938 (England): "The jurisdiction of this Court … as representative of the Crown, with regard to the custody of infants rests upon this ground, that it is the interest of the State and of the Sovereign that children should be properly brought up and educated; and … the Sovereign, as *parens patriae*, is bound to look to the maintenance and education … of all his subjects."

There are problems with this view. First, it makes *parens patriae* legally and analytically superfluous and its invocation therefore likely to confuse and mislead. More importantly, under the banner of *parens patriae*, the state has extended its decision-making authority to areas of persons' lives ordinarily treated as outside the proper ambit of state police power. These are areas ordinarily considered matters of self-determining right because they satisfy two conditions: (1) they concern central aspects of the right holder's own life, and (2) exercise of the right threatens no "harm" to others, even though it might affect others (e.g., by offending or upsetting them). Let us call these *Private Matters*, giving 'private' here the special meaning of having these two features. They include choices such as with whom one forms close personal relationships, whether and how one receives medical treatment for illness or injury, how one receives education, and who will provide other services relating to one's person or property. Any such choice might disappoint or anger other individuals or undermine the aims of some group.[46] Yet the state ordinarily does not presume, as part of its police power function, to decide these things for private parties nor to force private parties to take into account interests of other individuals, of any groups, or of society as a whole when making their own choices about such matters among available options.

Thus, to treat *parens patriae* as simply police power applied to nonautonomous persons' lives, the legal system presumably should either (1) constrict *parens patriae* action to the established normal bounds of police power action, deeming state involvement in Private Matters altogether improper even as to nonautonomous persons or (2) provide normative justification for treating as appropriate for police power action, such that everyone's interests weigh in state decisions, aspects of dependent person's lives comparable or equivalent to aspects of autonomous persons' lives deemed inappropriate for police power control (Private Matters). Doing 1 seems unsatisfactory because it leaves nonautonomous persons vulnerable in important areas of their lives. To fulfill 2, it would not suffice simply to point out that nonautonomous persons are unable to make certain decisions for themselves. For as described above, there is another available approach to making decisions for such persons—that is, the fiduciary conception of *parens patriae*.

To illustrate: if I, an autonomous person, have diabetes, whether I take insulin is a Private Matter; it concerns a central aspect of my life, and my decision in and of itself threatens no harm to others.[47] The legal system imbues me with

46 See Banks, *Is Marriage for White People?* 136–66, describing social pressure in the Black community not to "marry out."

47 Driving after choosing not to take insulin is a different story; traffic safety is not a Private Matter.

the right to decide; whether I take insulin is entirely up to me. A liberal state would not presume to compel me to do so on the grounds that it would be best for other individuals, such as my family members, or for society generally, even though that might be true. If *parens patriae* were not a different function relative to police power but rather just a special application of police power to situations in which nonautonomous persons are involved, then the legal system should either (1) cease exerting *parens patriae* authority in the realm of nonautonomous persons' medical care or (2) justify making those persons' medical care a proper subject of police power, to be dictated by the state's balancing of all affected interests such that, for example, the state might prohibit doctors from treating a particular diabetic child if it happened to be factually the case that her parents' interests in preventing treatment (e.g., because it conflicts with their religious beliefs) combined with any other affected private interests (e.g., of co-religionists, siblings in an overcrowded house, health insurance companies) and collective societal aims (e.g., avoiding costs of enforcing treatment mandates over parental objection, lowering medical care prices by reducing demand) outweigh the child's interests. No theorists have provided the justification called for by option 2. Section 5 will consider what justification the state has for exerting power over such aspects of nonautonomous persons' lives and whether it can support a police power approach.

Now consider the opposite position. The *parens patriae* role is completely distinct from the police power role. There is no overlap; they never pertain to the same situation. *Parens patriae* takes over where the police power role must stop. The state wears one hat when fulfilling police power duties and a different hat as *parens patriae*; it compartmentalizes these different roles and shifts from one to the other across contexts, just as a lawyer might represent an organization in a contract dispute today and represent an individual who happens to be a member of that organization in an unrelated criminal prosecution tomorrow. Police power extends to state efforts to improve general welfare and to prevent some persons—whether autonomous or not—from harming others. Call those *Public Matters*. They include some state action upon nonautonomous persons, such as civil commitment and juvenile delinquency proceedings aimed at protecting others in the community from harms a mentally ill adult or youth might cause. But on this position, police power does not extend to Private Matters for any persons, whether autonomous or not. Instead, *parens patriae* pertains to Private Matters for nonautonomous persons, and exclusively so; the state completely shifts focus from the collective to the individual. For present purposes, we need not establish where exactly the line lies between Public Matters and Private Matters.

Presumably, this differentiation between two decision realms, to correspond with treating *parens patriae* as entirely distinct from police power, would mean

the practical operation of the state's roles in the two realms differs. As noted, the police power role of government, as agent for all of society, is carried out by considering interests of all members of society, weighting objectively and balancing as necessary.[48] An operationally different *parens patriae* state function would therefore presumably entail the state considering interests of less than all members of society in rendering decisions within the purview of that function. Given that, by all accounts, this is a role adopted at least in part to protect in some fashion the welfare of nonautonomous persons, naturally, those nonautonomous persons whose lives are the subject of state decision would be included among those whose interests receive consideration. If they are the *only* persons whose interests may influence decisions, that would support adopting the fiduciary conception of *parens patriae*. The question then is whether, on a conception of *parens patriae* as distinct from police power, there are *any* other persons—or perhaps even groups of persons collectively—with interests impacted by decisions concerning Private Matters in nonautonomous persons' lives who ought also to be considered in *parens patriae* decision-making.

The answer cannot be *all* such other persons. That would return us to the position of no practical distinction between *parens patriae* and police power. The point of distinguishing the two roles must be to signal that with *parens patriae* decision-making, some who have interests at stake are not to be included within the state's scope of concern; their interests are real but irrelevant, properly disregarded. Now, if it can be shown that some subset of all other persons stands in a privileged position, such that they rightly have their interests considered even though the actually-impacted interests of some other persons are not, then we might settle on a third conception of *parens patriae*. Call it the *subgroup conception*: the state is not agent for all of society nor agent solely for nonautonomous individuals but rather serves some subset of society, of which an nonautonomous person is just one member. Note that the point of adding others' interests must be that these might conflict with and to some degree override the welfare of a nonautonomous person. If they were presumed entirely consistent with the latter's well-being, then it would be analytically superfluous to add them and to posit this third conception (except perhaps as a tiebreaker in the rare case when two alternatives are equally good for the nonautonomous person—a possibility discussed further below).

48 This is true even when the state is adjudicating competing individual rights and allegations of individual harm, as in contract and property disputes or enforcement of criminal laws. Legislatures and courts consider broader societal effects of potential decisions (e.g., general deterrence in connection with criminal law enforcement, public policy limitations on enforcement of terms in contracts, wills, and trusts).

As among all who might appear plausible candidates for this privileged position (spouses, siblings, offspring, parents, grandparents, other extended family members, neighbors, teachers, fellow town residents, religious or ethnic communities to which nonautonomous persons or their caregivers belong, etc.), it is not obvious on what objective basis one would distinguish among them. Or between any of them and individuals or groups who take an interest yet do not intuitively seem plausible candidates (e.g., right-to-life activists in connection with pregnancy-related decision-making for nonautonomous women, atheist bystanders in connection with regulation of religious schools). In other words, on what rational basis could one say "Well, *those* impacted persons should not be considered *at all* because this is *parens patriae* decision-making we are doing, not police power, but interests of *these* impacted persons *should* be considered and should be treated as a basis for potentially sacrificing the welfare of the nonautonomous person to some degree, even though this matter is equivalent to one that lies within the realm of self-determination for autonomous persons"—that is, is a Private Matter? Further analysis might reveal some plausible basis, but unless and until one is established we should—if *parens patriae* is to be a function distinct from police power—exclude *all* others. The equal moral status of nonautonomous persons gives rise to a presumption of equal protection, an equal moral right to have only their own interests considered in their Private Matters. That would leave us with the first, fiduciary conception of the state in its *parens patriae* role—that is, concerned solely for the welfare of the dependent individual.

A reaction many will likely have is that particular other individuals have a moral *right* in connection with such Private Matters in the lives of nonautonomous persons.[49] On that view, a state decision adverse to those other persons'

49 Scott Altman suggests a different basis for considering parental interests, but only indirectly ("Why Parents' Interests Matter"). He contends a fiduciary for children should do not what is best for children per se but rather what the child would choose if able—i.e., a substituted judgment, taking into account every consideration the child might if he or she were autonomous. This would include, Altman says, the child's love and gratitude toward parents. He notes in support that fiduciaries for incompetent adults are sometimes permitted to act on values and affections the ward displayed while competent—for example, in giving gifts. The problems with this idea are too numerous to present here, but to note a few: that children love and are grateful to their parents is far too broad a concept to guide state decision-makers; one would want actual evidence that at least most people after reaching an age when their views are independent and well informed retrospectively judge that their own welfare was or could have been justifiably sacrificed for the sake of their parents to a particular degree in particular circumstances. Such evidence does not exist, and it seems especially unlikely to be found with respect to state decisions about private matters; filial love and gratitude presumably are weakest toward birth parents who were so little capable or motivated when the child was born that a best-interest assessment of their parentage

wishes might in fact cause "harm to others," if the concept of harm includes any infringement of persons' supposed other-regarding moral rights. The state thus should instead aim in its decisions about nonautonomous persons' lives to avoid conflict with those other persons' wishes.[50] With respect to an autonomous person, the law today in liberal societies does *not* recognize rights of anyone else in connection with these matters, even though other individuals and groups might take an intense interest (e.g., a husband regarding a wife's use of contraception, Native American tribes regarding members' marital and residential choices), make great sacrifices for that person, or hold religious beliefs that ascribe such rights to them. But maybe something distinctive about being nonautonomous means it is appropriate to ascribe to others entitlements as to those aspects of one's life. For example, an elderly incompetent person's offspring might anguish over the type of long-term care facility the parent will enter, perhaps believing the parent's only chance at eternal salvation requires ending life in a facility run by the religious denomination the offspring recently joined and believing they bear a moral duty of the highest order to ensure that outcome. The fiduciary model of deciding for the parent would render the offspring's anguish and conviction of no direct relevance because frustrating the offspring does not constitute harm to them. Yet how can the state be so indifferent to family members' anguish and convictions? Have they no right in this situation?

---

would have been a sufficiently close call for gratitude to make the difference, or when legal parents demand because of their religious beliefs power to deny their children education or medical care. Further, children have feelings for many people in their lives, so presumably the state should also take that into account, and interests of some people within the child's love universe might conflict with interests of others in that universe. (For example, one parent might wish to exclude or disempower the other, grandparents might disapprove of parents' choices, etc.) And on the other hand, those other people in the child's world presumably love the child in return and would not want the child's welfare sacrificed at all for their sake, and we might impute to them also a desire that the state decision-maker take into account the possibility that the parents' current desires regarding the child's life are not consistent with the child's "true desires" because distorted by misinformation, ideology they might later abandon, selfishness, etc. Perhaps these problems are among the reasons why the law actually limits altruism by guardians and holders of a power of attorney for an incompetent person to giving modest gifts that in no way adversely impact that person and only insofar as is consistent with that person's gift giving while competent, if they ever were. See *Matter of Hourihan* (2020) WL 5049128, at *4 (NJ Super. Ct. App. Div. Aug. 27, 2020); and *Davis v. Davis*, 298 Va. 157, 835 S.E.2d 888 (2019).

50 There is scholarly debate over the best understanding of harm in the harm principle. See Folland, "The Harm Principle and the Nature of Harm." Most theorists aim to distinguish it from mere offense or upset by interpreting it as a substantial setback to basic interests, like physical integrity and self-determining liberties. See Turner, "'Harm' and Mill's Harm Principle," 299–300.

Section 5 will discuss normative obstacles to ascribing such rights to others in Private Matters in nonautonomous persons' lives, but here note the conceptual implications of doing so: this response amounts in effect to denying that there are any aspects of nonautonomous persons' lives that are inherently Private Matters (i.e., condition 2 above would not pertain). Then, one must also say either:

1. All aspects of nonautonomous persons' lives are Public Matters, with state decision making about them properly influenced by *any* other persons' desires, though especially those who have a moral right to the state's direct concern for their wishes, and by collective societal aims, taking us back to the no-distinction position; or
2. There is a realm of nonautonomous persons' lives, coextensive with autonomous persons' Private Matters yet not treated as such for nonautonomous persons, where decisions are fundamentally and centrally "about" them yet as to which
    a. Some others (but not all who take an interest) can and do have moral rights, of a sort (other-regarding) that no one has in relation to autonomous persons, such that they can be "harmed" by effects (e.g., anguish) not considered harms when occasioned by decisions autonomous persons make about their own lives, which "harms" can be a basis for sacrificing a nonautonomous person's well-being to some degree, whereas
    b. Interests of all persons lacking such moral rights are entirely irrelevant to state decision-making even though it impacts them.

Theorists have not directly addressed the possibility of such a realm of life for anyone, where the subgroup conception of state decision-making would pertain. There has been an effort among philosophers to mount a convincing theoretical defense for ascribing other-regarding moral rights to just one set of other persons in connection with the lives of just one set of nonautonomous persons—namely, parental rights regarding minor offspring. That exertion thus far has not been successful.[51] A common deficiency of the various approaches to mounting such a defense is their failure to identify any plausible general principle that can serve as the major premise in a syllogism that begins with empirical observations about the experience or actions of parents and ends with a conclusion of moral entitlement to other-determining legal power on their part. But in addition, there has been no effort to explain why opening the

51 See Dwyer, "Deflating Parents' Rights," in which I identify flaws in various types of arguments.

door to *some* third parties (parents) does not let in *all* whose interests could be impacted (i.e., why we should prefer 2 over 1 above). Or at least some additional persons. Extended family members or even neighbors might care greatly about, for example, whether the state passes legislation requiring academic accountability of a child's evangelical Christian school, and some such persons might have provided more care for that child than the parents have. In other words, attempts at making the positive case for ascribing control rights to parents have not attempted to show that any other persons impacted by state decisions as to central aspects of a child's life should be excluded from consideration. (Note that showing the parent-child relationship is unique cannot suffice; every type of relationship is unique (that is what it means to be a type), and no normative implications follow from uniqueness per se.)

Moreover, there has been no philosophical attention devoted to third-party rights regarding incompetent adults—that is, to whether spouses, offspring, or guardians of such adults similarly possess a moral right to legally effective power over their lives—power that could entail sacrificing what the state deems in their best interests. (Existing law recognizes no such right.[52]) Thus, much theoretical work would need to be done to support the subgroup conception of the state's role in aspects of nonautonomous persons' lives that are Private Matters for autonomous persons.

Another possibility some might suggest is that police power and *parens patriae* are distinct roles, but the state can operate in both roles in any given situation, if both are appropriate. One can find judicial opinions and scholarly writings that invoke both in support of a particular conclusion, where the two align.[53] Acting in both roles simultaneously, however, is a conceptual

52 *Kicherer v. Kicherer*, 400 A.2d 1097, 1100–1 (MD 1979): "a court of equity assumes jurisdiction in guardianship matters to protect those who, because of illness or other disability, are unable to care for themselves. In reality the court is the guardian; an individual who is given that title is merely an agent or arm of that tribunal in carrying out its sacred responsibility. . . . See generally Blackstone, *Commentaries*, 463. Whereas consanguinity is a factor that may well be given consideration by the chancellor in the appointment of a guardian because nearest of kin are more likely to treat a ward with kindness and affection . . . , appointment to that position rests solely in the discretion of the equity court. . . ."

53 See, e.g., *Prince v. Massachusetts*, 321 U.S. 158, 166 (1944): "Acting to guard the general interest in [a] youth's well-being, the state as *parens patriae* may restrict the parent's control. . . ." Also, "A democratic society rests, for its continuance, upon the healthy, well-rounded growth of young people into full maturity as citizens. . . . It may secure this against impeding restraints and dangers. . . . It is too late now to doubt that legislation appropriately designed to reach such evils is within the state's police power" (168–69). See also Weithorn and Reiss, "Providing Adolescents with Independent and Confidential Access to Childhood Vaccines," 799–800 ("Two of the best-known regulatory structures justified by both the *parens patriae* and police power authorities are compulsory education laws

impossibility if they are distinct. We have established that if *parens patriae* is distinct from police power, then it must be so because it requires excluding consideration of some persons' interests. The state cannot so exclude some persons' interests and still be doing police power decision-making because the latter entails considering everyone's interests. If *parens patriae* is distinct, the two roles are logically incompatible.[54] Could they operate sequentially? In theory, yes: state actors could begin their decision-making process by considering only the interests of a nonautonomous person (or of that person and some limited number of others, under the subgroup conception) and break any ties (i.e., outcomes equally good from that initial perspective) by considering what best serves aggregate societal welfare. That would likely be an extremely small set of cases; typically, from a *parens patriae* perspective, there is a clear ranking of options that are sufficiently different to generate strong preferences among third parties. And in practice, legal systems generally preclude fiduciaries from letting ulterior interests serve as a tiebreaker. They do so in part because of a human tendency to rationalize self-serving decisions that sacrifice others' well-being and in part because doing so implicitly sanctions instrumental use of vulnerable persons to benefit others, a normative problem (addressed further below) with taking a police power approach to state control of any persons' Private Matters.[55]

We are left, then, with three contenders for the best conception of the *parens patriae* role. On one, it is not distinct from the police power role; it is simply what we call police power when its exertion impacts nonautonomous persons—or more narrowly, when it is exerted in areas of nonautonomous persons' lives equivalent to aspects of autonomous persons' lives generally considered outside the proper ambit of police power authority (Private Matters)—that is, when application of police power is extraordinary. On this conception, we need justification for extending the police power function to those aspects of only some persons' lives, and arguably we should jettison the concept of *parens patriae* as misleading analytical surplusage. On the other conceptions, *parens patriae* is meaningfully different from police power and operates when

and prohibitions on child labor"); and Donnelly, "Inherent Jurisdiction and Inherent Powers of Irish Courts," 2135 ("where the parents for physical or moral reasons fail in their duty towards their children, the State as guardian of the common good, by appropriate means shall endeavour to supply the place of the parents, but always with due regard for the natural and imprescriptible rights of the child").

54 Harvard Law Review Association, "Developments in the Law," 1200: "Given the different premises and purposes of the police power and the *parens patriae* power, courts should apply different principles when they analyze laws based on these two powers."

55 See *Re Eve* [1986] 2 SCR 388 (Canada) § 82, stating that in making surrogate medical decisions for an incompetent adult, "a court … must exercise great caution to avoid being misled by this all too human mixture of emotions and motives."

police power authority is inapposite. This is when decisions to be made are of a type that autonomous persons are entitled to make for themselves on entirely self-regarding grounds. In one variant (the fiduciary conception), the state in the *parens patriae* role should concern itself *solely* with the welfare of the nonautonomous person in question. On another (subgroup) conception, it may aim to satisfy wishes or further interests of some other but not all interested persons.

Choice among the three options (special heed/no distinction; fiduciary model; subgroup) thus raises the question whether there is normative justification for the state's considering interests of any other persons when it makes decisions about central aspects of nonautonomous persons' lives under the *parens patriae* banner. If no, the fiduciary conception is best. If yes, choice between the special heed/no distinction and subgroup conceptions turns on whether the state should distinguish among persons other than an nonautonomous person in choosing whose interests properly influence its decisions about nonautonomous persons' Private Matters, treating interests of some as relevant but interests of others as irrelevant. Identifying the best conception of the *parens patriae* role seems therefore to rest on questions of justification for letting third-party interests influence state decision-making about nonautonomous persons' Private Matters and thus to require endorsing one or more normative premises, as well as narrowing the options on conceptual grounds.

5. Normative Bases for Selecting a Conception of *Parens Patriae*

Whether the state is justified in allowing other persons' interests to influence decision-making about Private Matters in nonautonomous persons' lives might depend on why the state ought to forbear from constraining, for the sake of other persons' interests or collective aims, autonomous persons' self-determining choices in those aspects of life. That reason might be inapplicable to nonautonomous persons, entirely or in certain contexts. Justification might instead or also depend on what affirmative warrant the state has for injecting itself into such aspects of life with nonautonomous persons.

As to the first possibility, there is of course a variety of philosophical accounts and political views as to where the limits of state power should be drawn in the standard case—that is, state interaction with autonomous persons. Space does not permit canvassing them all, let alone adjudicating among them. This section aims to establish simply that choice among conceptions of *parens patriae* could depend on to which basic normative outlook one generally adheres, assuming one is principled and aims for rational consistency across contexts and persons. The familiar exercise of contrasting deontological and utilitarian outlooks suffices to make the point.

A deontological view in fact predominates in international human rights discourse and in domestic individual rights jurisprudence and scholarship in Western liberal societies. In this outlook, the state owes *every* individual a presumptive negative duty of noninterference in private life because each has an inherent dignity that gives rise to rights of integrity and sovereignty over their own life and person.[56] For autonomous persons, that duty is overridden when individual choices threaten incursion on the integrity or liberty of others, otherwise not.[57] The inherent dignity is also incompatible with treating anyone as an object of other individuals' or any group's rights.

Declarations of rights emanating from this perspective rarely explicate the basis for ascribing dignity and therefore for the state's presumptive negative duty—that is, what it is about human individuals that commands respect. If presently existing autonomy were the sole source of moral worth and rights, the negative duty of noninterference might not be owed at all to nonautonomous persons, and they might be proper objects of anyone else's rights. This would leave the state free to insert itself into their lives however it wishes (unless it owes a duty of restraint regarding them to third parties) or to serve wishes of private parties who take a particular interest. Such a view of moral worth has little support among theorists today.[58] It is also contrary to prevailing moral intuitions reflected in the international human rights regime and Anglo-American legal systems; they ascribe right-conferring dignity to both autonomous and nonautonomous humans, as evidenced by the UN Conventions on children's rights and rights of persons with disabilities (including intellectual disabilities).[59] When rights declarations offer explanation, they typically include

56 Etinson, "What's So Special About Human Dignity?"; Bayefsky, "Dignity, Honor, and Human Rights"; and Kateb, *Human Dignity*. See also the preamble of the International Covenant on Civil and Political Rights: "these rights derive from the inherent dignity of the human person."

57 Folland, "The Harm Principle and the Nature of Harm"; and *Lawrence v. Texas*, 539 U.S. 558, 578 (2003): "The present case … does not involve persons who might be injured.… The Texas statute furthers no legitimate state interest which can justify its intrusion into the personal and private life of the individual."

58 See Clarke and Savulescu, "Rethinking Our Assumptions About Moral Status"; and Dwyer, *Moral Status and Human Life*, ch. 3. Typically, theorists operating from a deontological perspective, if they discuss nonautonomous humans at all, simply assume that they are persons with moral status equal to that of autonomous persons.

59 UNCRC, Preamble, which invokes "the inherent dignity and the equal and inalienable rights of all members of the human family"; and the UN Convention on Rights of Persons with Disabilities (UNCRPD), Preamble, which is similar. All major English-speaking nations other than the United States are parties to the UNCRC and the UNCRPD. The US is a signatory but not yet a party to either convention. The US Supreme Court has yet to recognize fundamental rights of young children, but it has rendered momentous decisions based on

reference to other things about humans that have currency among theorists of moral status, such as sentience, being subject-of-a-life, or capacity for development and flourishing.[60] Court decisions in many jurisdictions reflect this broader view of which persons have inherent dignity and a presumptive right of inviolability and noninstrumentalization, demanding (1) special justification for state exertion of power over intimate aspects of the lives of children or incompetent adults as well as of autonomous persons—specifically, justification tied to their welfare—and also (2) that the exertion not go beyond what that justification supports.[61] This implicitly rules out extension of police power into nonautonomous persons' Private Matters—or indeed, exerting state power over those aspects of their lives so as to serve *any* other persons' interests—thus suggesting *parens patriae* action in that realm is distinct and fiduciary in nature.

From a utilitarian perspective, on the other hand, one might say the state, at a metalevel of lawmaking, always properly considers the interests of all members of society, *à la* the police power, even regarding intimate aspects of autonomous persons' lives. It refrains in practice from substituting its own choices for those of autonomous persons in Private Matters only because it assumes that as to these areas of personal life, maximum aggregate societal welfare is generally optimized by ascribing legal "rights" of self-determination to such individuals. As per Mill, government control of those areas of life would not optimally promote societal welfare because, *inter alia*, (1) the individual has the greatest interests at stake and is better positioned than state actors to determine what best serves those interests, and (2) even when autonomous persons make poor choices, they and others learn from this and become more capable of

---

equality rights of young children, such as *Brown v. Board of Education of Topeka*, 347 U.S. 483 (1954), and it has ascribed several constitutional liberties to older children. Numerous lower courts have ascribed fundamental rights to young children, including protections of attachment relationships and bodily integrity.

60 For example, the preamble of the Universal Declaration of Human Rights cites "freedom from fear and want" and "social progress and better standards of life."

61 *H v. AC* [2024] NSWSC 40 (New South Wales) § 64: "Differing measures of legal protection are required according to the physical and mental capacities of individuals at particular times. Human dignity is a value common to municipal law and to international instruments relating to human rights. . . . Human dignity requires that the whole personality be respected: the right to physical integrity is a condition of human dignity, but the gravity of any invasion of physical integrity depends on its effect not only on the body but also upon the mind and on self-perception." See also *Re Kara* [2020] NSWSC 1083 (2020), § 65: "in exceptional cases where deprivation of liberty is a necessary consequence of the exercise of the *parens patriae* jurisdiction for the protection of the child and the promotion of his or her welfare, the making of orders by the Court as *parens patriae* that interfere with the personal integrity and liberty of a child will not contravene the child's human rights."

utility-maximizing decisions in the future.[62] From this outlook, at least part of the state's rationale for withholding the police power from Private Matters (or appearing to) in practice does not apply, might apply to a lesser degree, or might apply in some decision contexts but not others in cases involving nonautonomous persons. Though they also typically have the greatest interests at stake in those aspects of their lives, nonautonomous persons are by definition not in the best position to judge their own welfare, and the costs for them and others of bad decisions might outweigh the benefit of any learning they realize from their own mistakes.

On this consequentialist view of normal limits on police power, the government's *parens patriae* role might not be distinct. Reference to it might serve simply to remind decision makers that nonautonomous persons' interests matter and are weighty, per the special heed conception. Or to signal why state involvement is appropriate in those contexts for these persons. Or it might mark out a subcategory of decisions as to which state actors should focus exclusively on nonautonomous persons' interest, but for pragmatic rather than normative reasons; even a utilitarian view could support the fiduciary conception in practice in some decision contexts while allowing for the special heed conception to apply in others. (It is difficult to imagine how it could yield the subgroup conception in any context.) A utilitarian approach to regulation of nonautonomous persons' private lives thus appears more indeterminate than the deontological view in terms of which decision-making model is best in actual operation. Within it, any invocation of *parens patriae* would not reflect an inherent, normative constraint on normal police power state decision-making. Balancing of all affected interests would be morally appropriate, and if acting against nonautonomous persons' interests or wishes in some instances in Private Matters in their lives would promote aggregate welfare, the state should do that. For example, it might refuse to allow termination of an intellectually disabled woman's life-threatening pregnancy because it thinks the future child will be cognitively high functioning, and saving the child would satisfy preferences of family members and antiabortion activists.

Such an example makes doubtful that widespread commitment to utilitarian thinking best explains liberal states' current practices regarding Private Matters for anyone. One will be accused of moral vacuity if arguing for a position regarding such moralized issues as abortion or treatment of gender dysphoria in minors in terms of utility-maximizing cost-benefit analysis. One might further

62 See Mill, *On Liberty*; and *H v. AC* [2024] NSWSC 40, § 81: "competent adults are assumed to be 'the best arbiter[s] of [their] own moral destiny' and so are entitled to independently assess and determine their own best interests, regardless of whether others would agree when evaluating the choice from an objective standpoint."

test this sociological hypothesis thus: identify aspects of private life with respect to which that outlook's broad empirical assumptions regarding *autonomous* persons are doubtful (perhaps marital choice); then consider the likely popular reaction to a proposal that the state assume greater control of those. The point here is not to challenge any version of consequentialism as a political theory, which of course cannot be done by a counting of hands, but simply to identify for anyone drawn to that outlook its potential implications not only for *parens patriae* but also for treatment of autonomous persons. People can be pluralist in their ethical outlooks, applying consequentialist reasoning in some policy contexts and a deontological (or other) normative framework in other contexts, but it would be odd and require defense to switch normative frameworks—for example, from deontological to consequentialist—within the same area of life simply because a different group of human beings is under discussion.

Moving to the affirmative case for state involvement in Private Matters for nonautonomous persons, we should ask what moral basis the state has in the first place for exerting control over central aspects of their lives. The deontological view at least demands such justification, and modern social contract theories with deontological underpinnings make this burden of justification explicit.[63]

The state's affirmative justification for assuming control in some fashion of Private Matters for nonautonomous persons is straightforward. They have important needs they cannot satisfy themselves, and the state is best positioned to make certain decisions on their behalf, whether by making final choices directly or by choosing private surrogates to do so. As expressed by England's High Court of Chancery in 1827, the power to intervene in the family life of children "belongs to the King as *parens patriae*, having the care of those who are not able to take care of themselves, and is founded on the obvious necessity that the law should place somewhere the care of individuals who cannot take care of themselves, particularly in cases where it is clear that some care should be thrown round them."[64] Nonautonomous persons of any age need to enter into nurturing relationships with protected legal status, to be removed from relationships that prove damaging, and to be ensured daily care and treatment when ill or injured. Children need education, and many adults with intellectual disabilities

63 See Scanlon, *What We Owe to Each Other*; and Rawls, *Political Liberalism*, ix.

64 *Wellesley v. Beaufort*, 38 Eng. Rep. 236. See also *Durham Children's Aid Society v. BP*, O.J. No. 4183, § 29 (2007) (Ontario); *Re Eve* [1986] 2 SCR 388 (Canada) § 73 ("The *parens patriae* jurisdiction is … founded on necessity, namely the need to act for the protection of those who cannot care for themselves"); *Re F.* [1990] 2 AC 1 (invoking the common law "doctrine of necessity" to justify medical treatment of an incapacitated adult without consent); and Hall, "The Vulnerability Jurisdiction, 191 ("Simply put, the discretion is to do what is necessary for the protection of the person for whose benefit it is exercised").

can have much more flourishing lives if provided habilitation services. The state therefore aims to ensure these things for nonautonomous persons. Its doing so is supported by—rather than an insult to—the foundational value of human dignity. State respect for nonautonomous persons actually entails breaching the metaphorical wall of integrity around them to address their needs and, when possible, facilitate their maintaining or developing toward autonomy.[65] This seems a complete justification and the best, perhaps sole, one available from a deontological perspective. From a contractarian standpoint, we might characterize the warrant for state action in at least some nonautonomous persons' Private Matters as resting on their hypothetical consent. If able to make a rational choice, they would agree to state intrusion into their lives for this purpose.[66] Indeed, there is arguably actual consent in the case of those who have transitioned from autonomous to nonautonomous and while in the former condition endorsed their state's legal rules for treatment of them in the latter condition.

A final step in the analysis, enabling us to select between the fiduciary and subgroup conceptions of *parens patriae*, if operating from the deontological outlook, would be to recognize that because of the background universal right against intrusion into private life, the power the state exerts in nonautonomous persons' Private Matters must not outrun its justification.[67] A *least restrictive means* condition applies: intrude only so far as necessary to serve the aim that warrants your intruding at all.[68] That the state has *some* warrant for assuming control to *some* degree over nonautonomous persons' private life does not license the state to then use its control for any and all purposes it might choose, including treating these persons instrumentally to serve interests of other persons or collective entities. That would be contrary to respect.

65 Schapiro, "What Is a Child?"

66 Vallentyne, "Libertarian Perspectives on Paternalism," 182–93 (reserving the notion of hypothetical consent to persons who have had some degree of autonomy). We might imagine representatives of future persons in Rawls's "original position" agreeing to such a regime.

67 Smith, "Parenthood Is a Fiduciary Relationship," 428: "When a person, the fiduciary, acquires powers not for her own benefit but, rather, to allow her to attend to the interests of her beneficiary, then the fiduciary must use those powers in what she perceives to be the best interests of the beneficiary. That is what the powers are for; that is the basis on which and the purpose for which they are acquired; and that is how they must be used. Any other use is a misuse." See also *Re Eve* [1986] 2 SCR 388 (Canada) § 77: "It must be exercised in accordance with its underlying principle. . . . The discretion is to be exercised for the benefit of that person, not for that of others."

68 See Department of Health (UK), "Mental Health Act 1983," 23; and the webpage "Guardianship: Less Restrictive Options" from the Elder Justice Initiative (US Department of Justice), https://www.justice.gov/elderjustice/guardianship-less-restrictive-options (updated September 30, 2024; accessed February 28, 2025).

Imagining ourselves one day losing our self-governing faculties lends intuitive support to this conclusion. We would not now assent to the proposition that such an eventuality would create a legitimate opportunity for the state, without prior authorization by us while competent, to use us instrumentally in ways currently (while we are competent) impermissible without our consent—for example, to subject us to medical experimentation or organ harvesting or to empower a particular family member to dictate our treatment solely in order to placate that family member or to enable them to act on a religious command they believe themselves under. We would not assent to that even if we recognize that we will in that situation need the state to assume some authority over our personal lives for the sake of our own basic welfare. Nor even if the proposition were qualified with a side constraint such as that the medical experimentation, organ harvesting, or guardian appointment must not cause us *grievous* harm, or that the benefit to others must clearly and substantially outweigh the cost to us.[69] We might or might not believe we have a positive entitlement to state solicitude for us in our vulnerable situation, but we would expect that *if* the state chooses to breach the normative wall around us in reaction to our loss of autonomy and to exert power over our person, it will do so only in order to effectuate what it reasonably deems beneficial for us or to carry out wishes we expressed while competent. From another angle, we would say that losing our mental faculties does not amount to forfeiting our personhood, human dignity, or right against the state's treating us instrumentally in connection with central aspects of our lives. Rather, we would say, "Leave us alone except insofar as you are going to try to benefit us or carry out our prior choices." The state must act solely for the purpose that is the *raison d'être* of its power in these personal aspects of life. Presumptively, we should say the same of persons who have not previously possessed autonomy.

69 See Smith, "Parenthood Is a Fiduciary Relationship," 442–43: "In 1934, when the Dionne quintuplets were four months old, the Ontario government took them from their parents. . . . They were put into a kind of zoo, which millions of people paid to visit. . . . At the time, this might have seemed justifiable to some." See also Payton, "The Concept of the Person in the *Parens Patriae* Jurisdiction over Previously Competent Persons": "Were it not for the fiduciary nature of this custody, which gives the ward rights against his custodians, the incompetent's disappearance as an empowered legal person would work a forfeiture exceeding any punishment imposed under the criminal law. . . . The fiduciary nature of the *parens patriae* jurisdiction over formerly competent incompetents therefore is critical to the legitimacy of the state's exercise of power over them, since the state . . . would otherwise in effect confiscate the body and property of an incompetent human being, on the sole ground of his incompetence" (617). Further, "the King became the servant, not the master, of persons whom he brought under his protection. The powers of the state over the incompetent are tolerable only if fiduciary in nature and if administered in good faith out of fiduciary motive" (641).

The deontological view thus appears, again, from this different angle, to point to the fiduciary conception of *parens patriae* as best among the three identified—indeed, as the only acceptable conception. It is improper, within that normative framework, for the state, in making decisions about Private Matters in nonautonomous persons' lives, to choose on its own to benefit *any* other persons, even though others take an interest in the situations and will be affected in some way by the decisions. The state acts solely as agent for the nonautonomous individuals whose lives are at issue, with a duty of undivided loyalty.

As emphasized above, this assessment of connections between normative outlooks and *parens patriae* is not exhaustive and does not aim to establish a "truth of the matter." Its conclusion is simply that from the particular normative outlook underwriting the widespread conviction among competent adults that the state should not intrude into our own Private Matters, as among the three conceptions identified as distinct from each other and coherent, one should endorse the fiduciary conception of the state's role in nonautonomous persons' Private Matters. From a utilitarian outlook, in contrast, one might endorse extension of police power into some or all Private Matters, but then one should be prepared to accept that there is no in-principle obstacle to doing that with respect to our lives as autonomous persons as well. Further, it would presumptively be apt as to all nonautonomous persons, not just children, and it would entail consideration of all third-party interests, not just those of parents or other caretakers (whose interests might easily be outweighed by broader societal interests in most instances). But a utilitarian outlook, depending on the version of it deployed, might be indeterminate; it could conceivably also recommend the fiduciary conception, at least as to some types of decisions.

Still other normative perspectives might yield one or another definite conclusion or might also be indeterminate. It seems unlikely that any perspective would point toward the subgroup conception. Regardless of outlook, one should apply any perspective's fundamental principles consistently across persons, absent demonstration that they should not apply to some, and avoid ad hoc assertion and *sui generis* thinking about any group of nonautonomous persons. The fiduciary conception of *parens patriae* appears most consistent with the regime of individual rights that autonomous adults in liberal Western societies have come to expect and demand for themselves.

6. IMPLICATIONS

If any legal system were unambiguously to adopt the fiduciary conception of *parens patriae*, its scope would be limited to Private Matters. Legal actors would not refer to *parens patriae* in connection with state efforts to prevent

nonautonomous persons from harming others, as in juvenile delinquency proceedings; police power would be appropriate in those cases. In Private Matters, police power would be inappropriate, and state actors, when appropriately exerting control, would do so solely based on a nonautonomous person's welfare. As to rights of a constitutional nature, the state would be bound by none in connection with its control of Private Matters except those of the nonautonomous person. In essence, the state would act as an agent or surrogate for those private individuals, stepping into those persons' shoes, not in its usual role as agent for society as a whole.[70] The state would owe duties solely to the nonautonomous persons in question, and those would include duties (1) not to assume and exert power over those persons' lives beyond what their needs justify and (2) not to misuse its power by deploying it to serve others and thereby treat the nonautonomous persons instrumentally.[71]

As noted, some people have an intuition that certain family members have special claims in connection with decision-making for nonautonomous persons. The intuition appears largely confined, though, to parents raising minor offspring. For reasons theorists have not explored, it finds little expression in connection with family members caring for or concerned about incompetent adults. As to the latter, there is general acceptance that family members have no *entitlement* regarding state appointment of guardians or direction of particular decisions such as receipt of medical treatment, no matter how intense family members' feelings are about such things, and courts have taken that view.[72] Yet in connection with child-rearing, a common view, at least in the United States during the past century, is that parents themselves have moral rights—specifically, a right of biological parents to the state's making them legal parents and a right of legal parents to the state's conferring extensive legal

70 See Blokhuis, "Whose Custody Is It Anyway?" 207, citing *Young v. Young*, 4 SCR 3 (1993): "when a Canadian court issues custody and access orders in disputes between former spouses and domestic partners, it is not 'state action' subject to Charter scrutiny." See also Brief Amicus Curiae of the Ohio Association of Juvenile Court Judges, in *Gault* 1966 WL 100788, 8: "It is the unquestioned right and imperative duty of every enlightened government, in its character of *parens patriae*, to protect and provide for the comfort and well-being of such of its citizens as, by reason of infancy are unable to take care of themselves … and all constitutional limitations must be so understood and construed as not to interfere with its proper and legitimate exercise."

71 Criddle, "Liberty in Loyalty," 995, explaining that the duty of loyalty guards against domination "by ensuring that a fiduciary's actions are legally required to track the terms of her mandate and the interests of her beneficiaries."

72 See note 33 above.

powers and privileges on them.[73] The analysis above of *parens patriae* presents another way of assessing those claims, which to be coherent must be predicated on parents' own interests and thus effectively call for deployment of either the special heed/no-distinction conception or the subgroup conception of *parens patriae*.[74] The fiduciary conception categorically rules out the state's considering parents' interests for their own sake when it decides on behalf of children such Private Matters as with whom they will have legal family relationships, as among available and willing persons, or what types of education are permissible or compulsory for them. It also rules out controlling those aspects of their lives to serve collectivist interests.

Looked at another way, defense of parental *rights* (to be chosen as legal parents or to have certain legal powers), as opposed to parental *authority* imbued solely for the sake of and only so far as warranted by children's welfare, seems to require rejecting deontological views as bases for defining the limits of state power generally, in favor of some other view. Then one should also accept the broader implications of such an alternative view both (1) for the law governing nonautonomous persons' lives, including the possibility that persons other than parents also have rights to control children's lives, or at least a right to have their interests factor into decision-making and (2) for state intrusion into the private lives of *autonomous* persons.

Alternatively, that defense might rest on a demonstration that state decisions concerning parentage and legal-parent authority are actually Public Matters because denying what biological or legal parents demand would constitute harm to them. But this would require allowing that some nonparents might also be able to allege harm from decisions made regarding children—for example, infertile couples who wish to raise children who are not their biological offspring or taxpayers who are forced to bear the costs of bad state parentage decisions or bad parental child-rearing decisions. And it would require according the same treatment to, or somehow distinguishing, state decisions concerning appointment and empowerment of guardians for incompetent adults, because in that context as well, a family member (offspring, parent, sibling) might claim to be "harmed" if denied the opportunity to serve and control. It would further

73 For a contrary view in the United Kingdom, see *R. v. Gyngall*, 2 QB 239, stating that *parens patriae* is "not a jurisdiction to determine rights as between a parent and a stranger, or as between a parent and a child. It was a paternal jurisdiction … in virtue of which the Chancery Court was put to act … in the place of a parent, and as if it were the parent of the child, thus superseding the natural guardianship of the parent. On the absence of parental-rights thinking in early America, see Shulman, *The Constitutional Parent*.

74 On the incoherence of arguments for parental rights (as distinct from parental authority) that appeal to children's interests, see note 5 above.

require articulating a definition of harm that one is prepared to deploy consistently across all one's ethical views, including views as to one's own moral rights in making self-determining choices, which might mean others can correctly claim to be harmed by your personal choices (e.g., as to religion or intimate partner) and so constrict your rights.

7. CONCLUSION

It disserves persons in need of special solicitude to continue using *parens patriae* loosely to refer to any form of state attention to nonautonomous persons. Clarification of the role is long overdue. This article has initiated a more rigorous examination of the practice. It has identified three distinct and coherent conceptions of this state function, and it has shown how choice among them depends on normative assumptions regarding limits of proper state involvement in private life generally. Further, it has shown that the fiduciary conception, in which the state is subject to a duty of undivided loyalty, is the only one consistent with the prevailing understanding of why autonomous persons have a right of self-determination in connection with intimate aspects of their lives. This demonstration incidentally suggests need to reorient philosophical debate over parental rights so that it begins with focus on the nature and limitation of the state's role when the state presumes to render decisions about intimate aspects of children's lives and, at the same time, so that it examines this through a broader lens that encompasses all nonautonomous persons and all others with interests at stake in how their lives go.

William & Mary
jgdwye@wm.edu

REFERENCES

Altman, Scott. "Why Parents' Interests Matter." *Ethics* 133, no. 2 (2022): 271–85.

Archard, David, and Suzanne Uniacke. "The Child's Right to a Voice." *Res Publica* 27 (2021): 521–36.

Arstein-Kerslake, Anna. *Restoring Voice to People with Cognitive Disabilities.* Cambridge University Press, 2017.

Banks, Ralph Richard. *Is Marriage for White People? How the African American Marriage Decline Affects Everyone*. Dutton, 2011.

Bayefsky, Rachel. "Dignity, Honor, and Human Rights: Kant's Perspective." *Political Theory* 41, no. 6 (2013): 809–37.

Bishop, Joel Prentiss. *Commentaries on the Law of Marriage and Divorce and Evidence in Matrimonial Suits*, vol. 7. Little, Brown, and Company, 1852.

Blokhuis, J. C. "Whose Custody Is It Anyway? 'Homeschooling' from a *Parens Patriae* Perspective." *Theory and Research in Education* 8, no. 2 (2010): 199–222.

Bogert, George G., George T. Bogert, and Amy Morris Hess. *The Law of Trusts and Trustees: A Treatise Covering the Law Relating to Trusts and Allied Subjects Affecting Trust Creation and Administration (with Forms)*. West, 2019.

Brady, Maureen E. "Turning Neighbors into Nuisances." *Harvard Law Review* 134, no. 5 (2021): 1609–82.

Brereton, Paul L. G. "The Origins and Evolution of the *Parens Patriae* Jurisdiction." Lecture, Sydney Law School, May 5, 2017. https://supremecourt.nsw.gov.au/documents/Publications/Speeches/2017-Speeches/Brereton_050517.pdf.

Clarke, Steve, and Julian Savulescu. "Rethinking Our Assumptions About Moral Status." In *Rethinking Moral Status*, edited by Steve Clarke, Hazem Zohny, and Julian Savulescu. Oxford University Press, 2021.

Criddle, Evan J. "Liberty in Loyalty: A Republican Theory of Fiduciary Law." *Texas Law Review* 95, no. 5 (2017): 993–1060.

Criddle, Evan J., Evan Fox-Decent, Andrew S. Gold, Sung Hui Kim, and Paul B. Miller, "Introduction: Fiduciary Government, Provenance, Promise, and Pitfalls." In Criddle, Fox-Decent, Gold, Kim, and Miller, *Fiduciary Government*.

———, eds. *Fiduciary Government*. Cambridge University Press, 2018.

Custer, Lawrence B. "The Origins of the Doctrine of *Parens Patriae*." *Emory Law Journal* 27 (1978): 195–208.

Department of Health (UK). "Mental Health Act 1983: Code of Practice." January 2015. https://www.gov.uk/government/publications/code-of-practice-mental-health-act-1983.

Donnelly, Joan. "Inherent Jurisdiction and Inherent Powers of Irish Courts." *Judicial Studies Institute Journal* 2 (2009): 122–61.

Dwyer, James G. "Deflating Parents' Rights." *Law and Philosophy* 40, no. 4 (2021): 387–418.

———. "Inter-Country Adoption and the Special Rights Fallacy." *University of Pennsylvania Journal of International Law* 35, no. 1 (2014): 189–267.

———. *Moral Status and Human Life: The Case for Children's Superiority*. Cambridge University Press, 2011.

Etinson, Adam. "What's So Special About Human Dignity?" *Philosophy and Public Affairs* 48, no. 4 (2020): 353–81.

Folland, Anna. "The Harm Principle and the Nature of Harm." *Utilitas* 34, no. 2 (2022): 139–53.

Gold, Andrew S., and Paul B. Miller. "Introduction." In *Philosophical Foundations of Fiduciary Law*, edited by Andrew S. Gold and Paul B. Miller. Oxford University Press, 2014.

Graham, Anthony. "*Parens Patriae*: Past, Present, and Future." *Family Court Review* 32, no. 2 (1994): 184–207.

Hall, Margaret Isabel. "The Vulnerability Jurisdiction: Equity, *Parens Patriae*, and the Inherent Jurisdiction of the Court." *Canadian Journal of Comparative and Contemporary Law* 2, no. 1 (2016): 185–226.

Harvard Law Review Association. "Developments in the Law: The Constitution and the Family." *Harvard Law Review* 93, no. 6 (1980): 1156–383.

Hatcher, Daniel. "Purpose vs. Power: *Parens Patriae* and Agency Self-Interest." *New Mexico Law Review* 42, no. 1 (2012): 159–202.

Holland, Robert Wolstenholme. *The Law Relating to the Child, Its Protection, Education, and Employment*. Pitman and Sons, 1914.

Howard, Dana. "Deciding for the Incompetent." In *Routledge Handbook of the Philosophy of Paternalism*, edited by Kalle Grill. Routledge, 2018.

Joseph, Rosara. "Inherent Jurisdiction and Inherent Powers in New Zealand." *Canterbury Law Review* 11, no. 2 (2005) 220–38.

Kalb, Johanna. "Gideon Incarcerated: Access to Counsel in Pretrial Detention." *University of California at Irvine Law Review* 9, no. 1 (2018): 101–40.

Kateb, George. *Human Dignity.* Harvard University Press, 2011.

Kindred, Kay P. "God Bless the Child: Poor Children, *Parens Patriae*, and a State Obligation to Provide Assistance." *Ohio State Law Journal* 57, no. 2 (1996) 519–42.

Kramer, Matthew H. "Refining the Interest Theory of Rights." *American Journal of Jurisprudence* 55, no. 1 (2010): 31–39.

Legarre, Santiago. "The Historical Background of the Police Power." *University of Pennsylvania Journal of Constitutional Law* 9, no. 3 (2007): 745–96.

Leib, Ethan J., David L. Ponet, and Michael Serota. "Mapping Public Fiduciary Relationships." In *Philosophical Foundations of Fiduciary Law*, edited by Andrew S. Gold and Paul B. Miller. Oxford University Press, 2014.

Merkel-Holguin, Lisa, John D. Fluke, and Richard D. Krugman. *National Systems of Child Protection: Understanding the International Variability and Context for Developing Policy and Practice*. Springer, 2020.

Mill, John Stuart. *On Liberty*. John W. Parker and Son, 1859.

Miller, Paul B. "Fiduciary Representation." In Criddle, Fox-Decent, Gold, Kim, and Miller, *Fiduciary Government*.

Payton, Sallyanne. "The Concept of the Person in the *Parens Patriae* Jurisdiction over Previously Competent Persons." *Journal of Medicine and Philosophy* 17, no. 6 (1992): 605–45.

Pomeroy, John Norton. *A Treatise on Equity Jurisprudence: As Administered in the United States of America*. Bancroft-Whitney, 1883.

Rawls, John. *Political Liberalism*. Harvard University Press, 1993.

Scanlon, T. M. *What We Owe to Each Other*. Harvard University Press, 1998.

Schapiro, Tamar. "What Is a Child?" *Ethics* 109, no. 4 (1999): 715–38.

Seymour, John. "*Parens Patriae* and Wardship Powers: Their Nature and Origins." *Oxford Journal of Legal Studies* 14, no. 2 (1994): 159–88.

Shulman, Jeffrey. *The Constitutional Parent: Rights, Responsibilities, and the Enfranchisement of the Child*. Yale University Press, 2014.

Smith, Lionel. "Parenthood Is a Fiduciary Relationship." *University of Toronto Law Journal* 70, no. 4 (2020): 395–452.

Sreenivasan, Gopal. "Duties and Their Direction." *Ethics* 120, no. 3 (2010): 465–94.

Thomas, Gregory. "Limitations on *Parens Patriae*: The State and the Parent/Child Relationship." *Journal of Contemporary Legal Issues* 16 (2007): 51–62.

Turner, Piers Norris. "'Harm' and Mill's Harm Principle." *Ethics* 124, no. 2 (2014): 299–326.

Vallentyne, Peter. "Libertarian Perspectives on Paternalism." In *The Routledge Handbook of the Philosophy of Paternalism*, edited by Kalle Grill and Jason Hanna. Routledge, 2018.

Volokh, Eugene. "Parent-Child Speech and Child Custody Speech Restrictions." *NYU Law Review* 81, no. 2 (2006) 631–733.

Weithorn, Lois A., and Dorit Rubinstein Reiss, "Providing Adolescents with Independent and Confidential Access to Childhood Vaccines: A Proposal to Lower the Age of Consent." *Connecticut Law Review* 52, no. 2 (2020): 771–861.

Journal of Ethics and Social Philosophy
VOL. 30 NO. 3 · MAY 2025
https://doi.org/10.26556/jesp.v30i3.4137

RIGHTS INFRINGEMENT, COMPENSATION, AND LUCK EGALITARIANISM

Jesse Spafford

Suppose that *C* has a Hohfeldian claim against *A* ϕ-ing, where this claim implies a correlative moral duty on *A*'s part to not ϕ.[1] Further, suppose that *A* infringes on *C*'s claim by ϕ-ing in a way that imposes a cost upon *C*. Many rights theorists hold that *A*, in virtue of her ϕ-ing, acquires a unique duty to (at least partially) compensate *C* for this cost (where this duty is *unique* in that no one else acquires a similar duty to compensate *C* for her loss).[2] Call this proposition the *compensation thesis*. To illustrate the compensation thesis, consider the following case (adapted from a case originally produced by Joel Feinberg):

> *Permissible Infringement*: Hiker is deep in the mountains when a surprise blizzard strikes. She knows that she will die without shelter but fortunately stumbles across Owner's unoccupied cabin. Hiker breaks a window to enter the cabin and thereby survives the storm.[3]

1 For Hohfeld's complete schema of incidents, see "Some Fundamental Legal Conceptions as Applied in Judicial Reasoning." For a helpful summary, see Wenar, "Rights." For a presentation of Hohfeldian incidents that treats them as the component parts of moral rights (as opposed to legal ones), see Thomson, *The Realm of Rights*, 37–78.

2 See Feinberg, "Voluntary Euthanasia and the Inalienable Right to Life," 102; Davis, "Rights, Permission, and Compensation," 381–84; Lomasky, "Compensation and the Bounds of Rights"; Skorupski, *The Domain of Reasons*, 310–13; and Jorgensen Bolinger, "Moral Risk and Communicating Consent," 186n18. There are also theorists such as Thomson (*The Realm of Rights*, 95) and Frederick ("*Pro Tanto* Versus Absolute Rights," 388) who endorse the thesis but posit a set of exceptional cases in which compensation need not be paid, including cases in which the agent blamelessly infringes on another's claim. By denying that blameless infringement implies a duty to compensate, these theorists sidestep the argument of this paper. That said, they also do not provide a principled reason for excepting blameless infringements from the general duty to compensate wronged parties. This paper can be understood as providing such a reason. One might also think that *A* does not have to compensate *C* for any costs caused by her infringement that are unforeseeable. Here again, the paper can be understood as developing a principled basis for this carve-out.

3 For the original version of the case, see Feinberg, "Voluntary Euthanasia and the Inalienable Right to Life," 102.

There is widespread agreement that Hiker acts permissibly in this case—and that she acts permissibly even on the assumption that Owner has moral property rights such that Hiker has a correlative duty not to use the cabin. However, there are rival explanations of why Hiker's act is permissible. According to *specificationists*, the content of a right often includes tacit exception clauses that limit the set of actions against which the right holder has a claim. In this case, a complete specification of Owner's property rights would assert (among other things) that Owner has a claim that Hiker not use-the-cabin-when-Hiker's-survival-does-not-depend-on-using-the-cabin.[4] Given that Hiker's survival does in fact depend on using the cabin, Owner's claim does not apply to Hiker's action. Thus, Hiker acts permissibly because Owner has no claim against her action. Alternatively, there is a second variety of specificationism that reaches this same result without positing that there are exception clauses built into the content of Owner's claim. On this approach, the existence of the claim itself is held to be conditional on certain states of affairs obtaining.[5] Specifically, Owner's claim against Hiker using her cabin obtains only if Hiker's survival does not depend on that use. Given that Hiker's survival does depend on her use of the cabin, her circumstances negate Owner's claim, with this negation rendering Hiker's action permissible.

For these purposes, it will be helpful to set specificationism aside and focus on the primary rival account of why Hiker acts permissibly in Permissible Infringement. According to this influential view, Owner does in fact have a claim against Hiker using her cabin in virtue of Owner's property rights. However, this does not imply that Hiker ought not use the cabin; rather, Owner's claim is merely a *pro tanto* moral consideration that is overridden by the moral importance of Hiker's survival.[6] That Owner has an overridden claim in this case is evidenced by the apparent "moral residue" that it leaves behind (where this residue does not accompany more pedestrian permissible actions): even though Hiker acts in a way that is permissible, all things considered, she now owes it to Owner to pay for the repair of the cabin's window. In other words, the *pro tanto* proponents' reason for thinking that Owner's claim was infringed is that Hiker acquires a new unique duty to compensate Owner for the costs imposed by her infringement. They thus presuppose the compensation thesis: in any case where some agent *A* infringes on some claimholder *C*'s claim, *A* acquires a unique duty to compensate *C* for at least some of the costs imposed.

4 For a defense of this approach, see Shafer-Landau, "Specifying Absolute Rights."

5 See Wellman, "On Conflicts Between Rights."

6 This view is most famously endorsed by Feinberg, "Voluntary Euthanasia and the Inalienable Right to Life," 102; and Thomson, *The Realm of Rights*.

While the compensation thesis is both popular and intuitively plausible (as illustrated by the Permissible Infringement case), this paper will present two arguments for rejecting it. First, section 1 argues that rights theorists face a trilemma when it comes to cases in which *A* blamelessly infringes on one of *C*'s claims by ϕ-ing. Section 2 considers—and rejects—the proposal that the trilemma can be resolved by positing that *C* does not in fact have a claim against *A* ϕ-ing in such cases. Section 3 then presents an alternative way of resolving the trilemma that has been recently advanced by Renée Jorgensen.[7] However, section 7 argues that this proposal is implausible, as its supporting argument rests on a false dilemma. Thus, the paper concludes that the best way to resolve the trilemma is to reject the compensation thesis. The intervening sections (sections 4, 5, and 6) present a second argument against the compensation thesis, namely, that it is incompatible with (a plausible interpretation) of luck egalitarianism. Given this incompatibility, any rights theorist who endorses the thesis will be yoking their position to the negation of (a plausible interpretation of) a popular theory of distributive justice. Finally, section 8 considers six quick arguments for the compensation thesis and finds all of them lacking.

1. A TRILEMMA FOR RIGHTS THEORISTS

The first argument against the compensation thesis is that it gives rise to a trilemma that is best resolved by rejecting said thesis. This trilemma arises when one considers cases in which an agent blamelessly wrongs someone. Consider, for example, the following pair of cases (the former adapted from a case presented by Tom Dougherty and the latter adapted from a case presented by Jeff McMahan).

> *Consent Evidence*: An unforeseeable glitch causes Claimholder's computer to send Aggressor an email from Claimholder's account reading "Please ϕ, I want you to ϕ, and you have my permission to ϕ." Aggressor then ϕs on the basis of this message, causing Claimholder to incur significant costs as a result.[8]

> *Forfeiture Evidence*: Claimholder has an identical twin who is a well-known mass murderer. Unbeknownst to Claimholder, her twin is on a killing spree, and Aggressor has seen the twin's picture on the news. By coincidence, Claimholder's car breaks down in Aggressor's neighborhood, and Claimholder goes up to Aggressor to ask for help. Before

7 See Jorgensen Bolinger, "Moral Risk and Communicating Consent."

8 See Dougherty, *The Scope of Consent*, 67.

> Claimholder has a chance to say anything, Aggressor knocks Claimholder unconscious by ϕ-ing. Claimholder incurs significant costs as a result.[9]

In both cases, Aggressor seems to blamelessly wrong Claimholder. She wrongs Claimholder because Claimholder has (by hypothesis) a claim against Aggressor ϕ-ing in both cases, and Aggressor wrongs Claimholder when she fails to discharge the correlative obligations. By contrast, the blameworthiness of an action depends (at least in part) on the agent's epistemic state, where this state might include her knowledge, her beliefs, the evidence she possesses, and/or the available evidence that she does not possess but has a duty to acquire, depending on one's particular theory of blameworthiness. For these purposes, the exact set of necessary and sufficient conditions of being blameworthy will be left unspecified, but it will be assumed that a person blamelessly ϕs if both her possessed evidence and the *relevant* evidence—i.e., the available evidence that she *should* have gathered (a notion that will also be left unspecified)—suggest that a state of affairs obtains wherein it would be permissible for her to ϕ.

When this placeholder account is applied to Consent Evidence and Forfeiture Evidence, one gets the result that Aggressor is blameless in both cases. In the former case, the evidence that Aggressor possesses suggests that Claimholder has consented to Aggressor ϕ-ing, where such consent would negate Claimholder's claim against Aggressor ϕ-ing.[10] Granted, there was available evidence that Aggressor could have gathered that would have suggested a different conclusion. For example, Aggressor could have called Claimholder to confirm that she has Claimholder's permission to ϕ, where this would have resulted in Claimholder clarifying that she does not in fact intend or want Aggressor to ϕ. However, it does not seem like Aggressor was epistemically negligent in forming her belief that Claimholder consented based on the email alone. Thus, her relevant evidence is coextensive with her possessed evidence in this case. Given that this evidence suggests that Claimholder has consented to Aggressor ϕ-ing—i.e., it is permissible for Aggressor to ϕ—it follows that Aggressor blamelessly ϕs in Consent Evidence.

Similar remarks apply to the Forfeiture Evidence case. There, the evidence possessed by Aggressor suggests that Claimholder poses an imminent threat to her life, where posing an imminent threat is typically taken to negate claims

9 See McMahan, *Killing in War*, 164.

10 Whether some piece of evidence suggests that Claimholder consents to Aggressor ϕ-ing depends on one's preferred theory of which acts qualify as consent. That said, the email that Aggressor receives in Consent Evidence would be evidence that Claimholder consents to ϕ-ing on practically any theory of consent.

against aggressive actions that would disarm that threat and are proportionate to the threat. Stipulating that Aggressor ϕ-ing would be proportionate to a threat on her life, it follows that Aggressor's possessed evidence suggests that ϕ-ing is permissible in her circumstances. And again, it does not seem like Aggressor was negligent in failing to gather additional evidence prior to ϕ-ing. Thus, her ϕ-ing is blameless, as both her possessed evidence and the relevant evidence suggest that ϕ-ing is permissible.

Granted, the proposed account of blameless action is speculative, and different theorists might fill in the details in different ways. However, any adequate theory will seemingly yield the same result—namely, that Aggressor acts blamelessly when she ϕs. Thus, in both Consent Evidence and Forfeiture Evidence, Aggressor wrongs Claimholder but does so blamelessly.

The trilemma for rights theorists arises when the compensation thesis is applied to these cases of blameless wronging. Recall from above that this thesis holds that if *A* wrongs *C*, then *A* alone must (at least partially) compensate *C* for the costs that *C* incurs in virtue of *A*'s ϕ-ing. Given that Aggressor wrongs Claimholder in both Consent Evidence and Forfeiture Evidence, the thesis implies that Aggressor must make substantial transfers to Claimholder such that Aggressor internalizes the costs of ϕ-ing rather than Claimholder. However, this seems unfair. After all, in both cases, Aggressor is morally fastidious and acting appropriately in light of her possessed evidence. Further, it does not seem that she has negligently failed to gather additional evidence relevant to assessing the permissibility of ϕ-ing. Together, these facts support the judgment that it would be unfair if Aggressor had to fully internalize the costs that ϕ-ing imposes on Claimholder.

If one thinks that any acceptable theory of rights must fairly distribute the costs of rights infringements, then a contradiction is reached.[11] It cannot simultaneously be the case that (1) Aggressor wrongs Claimholder in Consent Evidence/Forfeiture Evidence, (2) the compensation thesis is true, and (3) the unfairness of Aggressor internalizing the costs of ϕ-ing in Consent Evidence/Forfeiture Evidence disqualifies any theory that implies that Aggressor must do this. Thus, one of these independently plausible propositions must be rejected.

11 This premise is notably presupposed by Jorgensen, whose defense of the proposition that it is fair for Aggressor to internalize her imposed costs is discussed at length in sections 3 and 7 below (Jorgensen Bollinger, "Moral Risk and Communicating Consent"). Those writing on self-defense also tend to insist that fairness is a decisive consideration when determining who should internalize the costs of threatened rights infringements. See, e.g., Otsuka, "Killing the Innocent in Self-Defense"; and Gordon-Solmon, "What Makes a Person Liable to Defensive Harm?" For a discussion of how the literature on self-defense interacts with the argument of this paper, see notes 40 and 42 below.

2. THE NONINFRINGEMENT SOLUTION

One possible resolution of the trilemma is to reject proposition 1—i.e., deny that Aggressor wrongs Claimholder in either Consent Evidence or Forfeiture Evidence. Call this the *noninfringement solution*. One way to formulate this denial is to hold that Claimholder initially held a claim against Aggressor ϕ-ing but lost that claim, either because she consented to Aggressor ϕ-ing or acted in a way that caused her to forfeit her claim. Given that Claimholder no longer has a claim against Aggressor ϕ-ing, Aggressor does not wrong Claimholder and thus does not owe her any compensation.

However, this proposal seems like an unacceptable way of resolving the trilemma. As a matter of pretheoretical intuition, it seems clear that that Claimholder does not consent to Aggressor ϕ-ing in Consent Evidence (or in Forfeiture Evidence, where nothing even resembling consent occurs). Further, while there are many theories of what constitutes consent, none supports the judgment that Claimholder consents in this case. On some views, Claimholder consents to Aggressor ϕ-ing when Claimholder forms some positive attitude toward Aggressor ϕ-ing (e.g., she intends that Aggressor ϕ).[12] Alternatively, some hold that Claimholder consents to Aggressor ϕ-ing when Claimholder communicates some relevant bit of information to Aggressor (e.g., that Claimholder intends that Aggressor ϕ).[13] And others hold hybrid views that attach other necessary or sufficient conditions to the formation of the aforementioned attitudes and/or communicative acts. For example, Jorgensen posits that *either* a relevant attitude *or* an act of communication is sufficient for consent.[14] And Dougherty holds (roughly) that Claimholder consents iff she issues a relevant directive or permission to Aggressor *and* various evidentiary conditions obtain that would justify Aggressor's belief that Claimholder issued the directive/permission.[15]

Irrespective of which of these accounts one favors, one will not get the result that Claimholder consents to Aggressor ϕ-ing in Consent Evidence. While Aggressor has evidence that Claimholder intends that Aggressor ϕ and issued a directive that she ϕ, Claimholder does not in fact form any relevant pro-attitude toward Aggressor ϕ-ing and does not issue a directive or communicate any information at all to Claimholder. Given that all of the listed accounts of consent take either attitude formation or an attempted communicative act

12 See Hurd, "The Moral Magic of Consent"; and Alexander, "The Moral Magic of Consent (11)" and "The Ontology of Consent."

13 See McGregor, *Is It Rape?* 124; and Dougherty, "Yes Means Yes." Tadros holds that consent requires only an *attempt* to communicate ("Causation, Culpability, and Liability").

14 Jorgensen Bolinger, "Moral Risk and Communicating Consent."

15 Dougherty, *The Scope of Consent*.

to be a necessary condition of consent, it follows that Claimholder does not consent to Aggressor ϕ-ing. Granted, the foregoing list of accounts is just a quick survey, and there are many other accounts that have not been considered. However, it seems unlikely that any omitted account will hold that Claimholder consents in Consent Evidence for the simple reason that Claimholder does not act at all—indeed, it can be stipulated that her mental states do not even change—and seemingly any account of consent will hold that Claimholder must act or, at the very least, undergo some change in state if she is to consent to some action. Thus, it does not seem that Claimholder could have consented to Aggressor ϕ-ing in Consent Evidence (or Forfeiture Evidence).

Similar remarks apply to theories of forfeiture. Irrespective of which theory one endorses, one will seemingly hold that Claimholder must act or change in some way to forfeit a right. But given that, by hypothesis, she takes no action and undergoes no change in either Consent Evidence or Forfeiture Evidence, it does not seem that Claimholder could forfeit her claim against Aggressor ϕ-ing.[16]

Given that Claimholder neither consents to Aggressor ϕ-ing nor forfeits a claim against Aggressor ϕ-ing in either Consent Evidence or Forfeiture Evidence, the proponent of the noninfringement solution must maintain that Claimholder either loses her claim involuntarily or never had a claim against Aggressor ϕ-ing in the first place. With respect to the latter proposal, one might argue that a complete specification of Claimholder's claim in Consent Evidence would reveal that its content is actually qualified such that Claimholder has a claim against Aggressor ϕ-ing *except when* Aggressor receives an email from Claimholder's account telling her that she may ϕ. Similarly, in Forfeiture Evidence, one might hold that Claimholder has a claim against Aggressor ϕ-ing *except when* Aggressor is justified in believing that Claimholder poses an imminent threat to her life.

This specificationist approach is the alternative way of assessing the Permissible Infringement case discussed above. Recall that specificationists explain the all-things-considered permissibility of Hiker's actions by holding that Owner's claim against Hiker using her cabin is really a claim against Hiker using her cabin when Hiker's life is not otherwise at risk. However, applying this

16 Jonathan Quong argues that whether some person *G* has a claim against *A* ϕ-ing depends on various facts about *A*, including *A*'s beliefs, how costly it would be for *A* to not ϕ, whether she could have avoided the choice to ϕ versus incurring the costs of not ϕ-ing, etc. Does such a view imply that Claimholder might lose her claim—even absent any action by Claimholder or change of Claimholder's state—simply because one of these facts about Aggressor changes? Quong affirms that this is not the case: even if one accepts his view, Claimholder loses a claim only if she acts in certain relevant ways, as such a restriction is needed to both make his view cohere with core judgments about claim loss and ensure that claimholders have adequate control over which claims they possess ("Rights Against Harm," 262).

specificationist approach to Consent Evidence and Forfeiture Evidence seems misguided. The appeal of specificationism is that it can explain why it is permissible for Hiker to break into the cabin without making claims merely *pro tanto* considerations. But in Consent Evidence and Forfeiture Evidence, Aggressor's ϕ-ing is seemingly *not* permissible, all things considered. Thus, positing that Claimholder's claim in each case contains an exception clause renders the proposed theory of rights extensionally inadequate. By denying that Claimholder has a claim against Aggressor ϕ-ing, the specificationist version of the noninfringement solution avoids the implication that Aggressor must bear the unfair burden of compensating Claimholder; however, the absence of a claim against ϕ-ing also unacceptably implies that Aggressor acts permissibly when she ϕs.

Similar remarks apply to the proposal that Claimholder involuntarily loses her claim against Aggressor ϕ-ing in both Consent Evidence and Forfeiture Evidence. If Claimholder does in fact lose her claim in this way, then there is no remaining basis for declaring Aggressor's action to be impermissible, all things considered. Insofar as one wishes to preserve this all-things-considered deontic judgment, one must reject the involuntary claim loss version of the noninfringement solution as well.

3. THE COMPARATIVE FAIRNESS SOLUTION

The previous section has argued that there is no version of the noninfringement solution that can adequately resolve the trilemma of section 1. This leaves two other propositions as candidates for rejection: the compensation thesis and the proposition that it would be unfair for Aggressor to internalize the costs of ϕ-ing in Consent Evidence and Forfeiture Evidence by compensating Claimholder. This section will consider a defense of rejecting the latter proposition that has been raised by Jorgensen in her discussion of cases of blameless wronging.[17] Call this proposal—i.e., the contention that it is not disqualifying for a theory to imply that Aggressor must internalize the costs of ϕ-ing—the *comparative fairness solution*.

17 See Jorgensen Bolinger, "Moral Risk and Communicating Consent." Jorgensen advances this argument in the context of defending her account of consent described in section 2 above. Because she maintains that *P* can consent to *Q* ϕ-ing via communication even when *P* does not intend that *Q* form any particular belief, she has to posit an account of communication via *signals*—and, more specifically, signaling conventions that do not require intentions on the part of the signaling party (194–95). However, the general argument of hers discussed in this section does not rest upon this account of communication (or her account of consent more generally).

While Jorgensen concedes that there is something superficially troubling about a morally and epistemically fastidious Aggressor having to uniquely internalize the costs of ϕ-ing, she argues that this result should not be taken to be disqualifying, as one cannot simply assess the fairness/acceptability of Aggressor internalizing the costs of ϕ-ing in isolation; rather, the assessment must be *comparative* such that one considers the fairness/acceptability of the alternative(s) to having Aggressor internalize these costs. Thus, one must consider whether it would be fair/acceptable for Claimholder to internalize the costs of Aggressor ϕ-ing in cases of blameless wronging rather than Aggressor internalizing them. And Jorgensen plausibly contends that this would be even *more* unfair than Aggressor internalizing the costs.[18] In defense of this point, one might note that even though Aggressor was morally fastidious, the imposed costs are still attributable to her choice, while Claimholder had no such direct causal connection to the generation of costs. Given this lack of connection, it seems more plausible that Aggressor should internalize the costs, even if there is something theoretically unattractive about this outcome given Aggressor's blamelessness.

Additionally, Jorgensen points out that it would be even more unfair for the victims of rights infringements to have to internalize the associated costs in cases where some members of society are more prone to having their rights blamelessly infringed than others.[19] For example, suppose that there is a social practice that consistently generates misleading evidence about Claimholder's intentions such that multiple agents Aggressor and Infringer blamelessly infringe on Claimholder's claims by ϕ-ing. If each infringing agent has to internalize the costs that she respectively imposes, then the collective imposed costs in this case would be divided across Aggressor and Infringer. By contrast, if Claimholder has to internalize the costs of infringement, then all of those costs would be concentrated on her, which seems like a much less fair outcome.

In other words, irrespective of whether one endorses the compensation thesis, one must have a view about who should internalize the costs in cases of blameless infringement like Consent Evidence and Forfeiture Evidence. Given the comparative implausibility of making Claimholder internalize those costs, one should hold that Aggressor must internalize these costs. Thus, one should resolve the trilemma presented in section 1 by endorsing the comparative fairness solution: it is not disqualifying for a theory or thesis to imply that Aggressor internalizes the costs of ϕ-ing, as all theories must ultimately have

18 Jorgensen Bolinger, "Moral Risk and Communicating Consent," 202.

19 Jorgensen Bolinger, "Moral Risk and Communicating Consent," 203.

this implication to avoid the comparatively worse implication that Claimholder must internalize the costs of Aggressor ϕ-ing.

While this proposed solution to the trilemma is superficially plausible, it will subsequently be argued that it does not succeed. First, however, sections 4–6 will provide an independent argument for rejecting the compensation thesis—namely, that the thesis is incompatible with (a plausible interpretation of) luck egalitarianism. This means that proponents of the thesis must incur the theoretical cost of committing themselves to the rejection of a (plausible interpretation of a) popular theory of distributive justice. Section 7 will then return to the comparative fairness solution, arguing that the previously introduced luck egalitarian approach reveals that Jorgensen's argument for this solution rests on a false dilemma. If one accepts a luck egalitarian theory of the kind described below, then one will have a principled basis for distributing the costs of Aggressor's ϕ-ing in an alternative way that is more plausible than either of the two options discussed just above. Thus, one cannot defend the implication that Aggressor must internalize the costs of ϕ-ing on the grounds that it is the most plausible of all the available options.

4. THE COMPENSATION THESIS AS A RIGHTS-BASED ALLOCATIVE THESIS

Before completing the foregoing argument against the compensation thesis, a second argument against the thesis must be introduced. This argument contends that the compensation thesis should be rejected because it is incompatible with (a plausible interpretation of) an influential theory of distributive justice, namely, luck egalitarianism. However, to show that these two positions are incompatible, one must first show that they both pertain to the same subject matter. To this end, the present section and section 5 will demonstrate that both positions are *rights-based allocative* theories of justice in the sense that they both (a) assign persons rights where (b) that assignment is a function of a prior judgment about the appropriate share of advantage that at least some persons should possess (where 'advantage' is a placeholder term referring to whatever it is that matters morally when it comes to distributive justice).[20] Having demonstrated this, section 6 will then argue that the two positions are incompatible, as they imply contradicting rights assignments, with luck egalitarianism implying the existence of rights that are negated by the compensation thesis. Thus,

20 Note that a theory need not provide a complete ascription of all of the rights possessed by all persons to qualify as a rights-based allocative theory; rather, it might merely posit *some* rights on the basis of a prior judgment about appropriate shares of advantage, with complementary theories then positing additional rights to complete the picture of which rights persons possess.

endorsing the compensation thesis requires rejecting luck egalitarianism (and vice versa).

That the compensation thesis implies an assignment of rights is readily apparent, as it asserts that person *A* owes duties of compensation to person *C* when *A* infringes one of *C*'s claims. Additionally, it is an *allocative* thesis in the following stipulative sense: the particular rights that it assigns to *C* and correlative duties that it assigns to *A* are a function of the shares of advantage that *A* should distribute to herself and *C ceteris paribus*.[21] Note that the compensation thesis is a thesis about how imposed costs are to be distributed in cases of rights infringement, where any theory of cost distribution is ultimately a theory about how *A* is to distribute advantage: to say that *A* should incur a cost of *x* in virtue of ϕ-ing is just to say that *A* should give herself a share of advantage equal to however much advantage she ends up with after ϕ-ing minus *x*. Finally, note that the thesis's assignment of rights is a function of these prescribed shares, as it assigns rights in such a way that agents discharging their correlative duties would realize the prescribed distribution of advantage. For example, in Consent Evidence, the thesis assigns Aggressor a remedial duty of compensation such that Claimholder does not end up with an improper share of advantage. More specifically, it holds that Aggressor, in virtue of her blameless ϕ-ing, uniquely loses rights—specifically Hohfeldian permissions—such that Claimholder ends up with the same share of advantage that she would have had otherwise, with Aggressor thereby fully internalizing the costs imposed by

21 This is not to say that *A*'s duty to compensate *C* is grounded in the fact that *A* should distribute certain shares of advantage to *A* and *C*, *ceteris paribus*. Rather, the fact that *A* has a compensatory duty is grounded in her prior infringement of *C*'s claim. However, the *particular* duty of compensation that is owed *is* grounded in the shares of advantage that *A* should allocate to *A* and *C*. Suppose that *A* trespasses on *C*'s property, causing *C* significant emotional distress. Given this costly infringement, the compensation thesis would assign *A* a duty to compensate *C* by providing a payment of, say, $1,000. The fact that *A* now has a compensatory duty—as opposed to no duty at all—is explained by the fact that *A* infringed *C*'s claim. However, if the question is what explains why *A* owes *C* $1,000 as opposed to $5 or $2,000, then the answer is facts about what shares of advantage *A* and *C* should possess in light of *A*'s infringement. Note that the mere fact that *A* infringed *C*'s claim cannot explain why *A* owes $1,000 versus $5. Thus, some other fact must function as the *explanans* beyond the fact of *A*'s infringement. Specifically, most rights theorists would seemingly affirm that *A* owes $1,000 to *C* because *C* is entitled to the share of advantage that she would have possessed absent *A*'s infringement (perhaps excluding any loss of advantage that *A* could not have foreseen). That facts about appropriate shares play this grounding/explanatory role is what makes the compensation thesis an allocative thesis (with this explanatory relation being what makes the thesis about appropriate shares "prior" to the assigned rights), without there being any contradiction between this proposed grounding relation and the proposition that *A*'s compensatory duty is grounded in the fact that she infringed on *C*'s claim.

ϕ-ing. Or alternatively, Aggressor might uniquely lose claims against Claimholder acting in ways that would make Aggressor internalize those costs. For example, she might lose claims against Claimholder seizing some of her holdings. In this way, the compensation thesis is revealed to be a thesis that assigns and negates rights—both permissions and claims—such that the discharging of all corresponding duties would realize a particular distribution of advantage among at least certain persons.

One might object that in describing the compensation thesis as an allocative thesis—with all the associated talk of distributing appropriate shares of advantage—the foregoing argument conflates *corrective justice* and *distributive justice*.[22] This distinction is one often drawn by torts theorists who see torts as a matter of corrective justice and *not* distributive justice.[23] Most minimally, those who endorse this distinction hold that principles of distributive justice assign shares of advantage while principles of corrective justice apply only in circumstances where an agent has wrongfully interfered with another.[24] In other words, corrective principles are distinct from distributive ones in that the former apply to only a subset of cases to which the latter apply. Given that the compensation thesis applies only in this proper subset of cases, one might conclude that it qualifies as a corrective principle rather than as a distributive principle. Some theorists elaborate on this minimal account by suggesting that principles of corrective justice differ from distributive principles in that they posit a more limited set of persons who either owe—or are owed—advantage-producing actions. For example, Jules Coleman argues that principles of corrective justice imply that only the wrongdoer has a duty to deliver resources to the wronged party; by contrast, distributive principles imply that other agents (e.g., the state) have duties to ensure that each person ends up with her appropriate share of advantage.[25] Finally, some theorists contend that corrective principles are distinct from distributive principles because the former imply a distinct set of rival prescriptions about which transfers persons should make.[26] On this

22 I am indebted to an anonymous reviewer for raising this worry.

23 For some influential examples of corrective theories of torts, see Weinrib, *The Idea of Private Law*; Coleman, *The Practice of Principle*; Gardner, *Torts and Other Wrongs*; and Ripstein, *Private Wrongs*.

24 See Gardner, *Torts and Other Wrongs*, 40; and Miller, "Justice," sec. 2.2. Strictly speaking, Miller puts things in terms of interference with *legitimate holdings*. However, unless one endorses the self-ownership thesis, it is not clear that Aggressor is interfering with Claimholder's holdings in Consent Evidence and Forfeiture Evidence, despite these being paradigmatic circumstances of corrective justice.

25 Coleman, *Risks and Wrongs*, 310–11. See also Weinrib, *The Idea of Private Law*, 71.

26 See Miller, "Justice," sec. 2.2; and Goldberg and Zipursky, *Recognizing Wrongs*, 356.

account, principles of corrective justice imply that wrongdoers should make remedial transfers to their victims even when doing so upsets the distribution of advantage that is just according to the correct principle of distributive justice.

For these purposes, one can grant there is a distinction between distributive justice and corrective justice, as the incompatibility argument that will be made in section 6 requires affirming only that the compensation thesis is an allocative principle in the very weak sense described above: it must simply assign at least some person(s) a duty to ensure that another person ends up with a particular share of advantage. So long as the compensation thesis assigns duties in this way, it will have overlapping subject matter with the interpretation of luck egalitarianism that will be provided in section 5, thereby allowing for the subsequent demonstration that the two positions are incompatible.[27] Granted, the compensation thesis may not qualify as a distributive principle as the tort theorists above use the term. However, for the purposes of this argument, the thesis need not be distributive in this strong sense, and this is not what is asserted above. Rather, the contention advanced here is merely that the compensation thesis has at least some implications for when persons have duties to ensure that others attain a particular share of advantage, where this makes the thesis of a kind with luck egalitarianism (albeit a more general kind than what the tort theorists label "distributive principles").

5. Luck Egalitarianism as a Rights-Based Allocative Theory

To show that luck egalitarianism has the same subject matter as the compensation thesis—and that the two positions are in fact incompatible—the theory must be presented in some detail. Specifically, luck egalitarians hold that the distribution of advantage is just if and only if any inequality corresponds to some sanctionable choice on the part of the worse-off.[28] Different luck egalitar-

27 Indeed, despite holding that there is a difference in kind between corrective principles and distributive principles, Weinrib grants that the two have overlapping subject matter (*The Idea of Private Law*, 70).

28 This statement of luck egalitarianism must be precisified in two respects. First, 'corresponds' must be read as implying that the inequality is proportionate to the sanctionable choice in question (as a given sanctionable choice might justify an inequality of only a certain magnitude). Second, many luck egalitarians hold that there is an additional individually necessary and jointly sufficient condition of a distribution being just: not only must any inequality imply a corresponding sanctionable choice, but any sanctionable choice must also imply a corresponding inequality. That said, some luck egalitarians such as Shlomi Segall reject this additional necessary condition (*Why Inequality Matters*). Additionally, note that 'sanctionable' here should not be read as implying some sort of wrongdoing on the part of the worse-off party. Rather, a sanctionable choice is any choice

ian theories then render this proposition fully determinate by specifying what advantage includes and which choices count as "sanctionable."

One can get a pretheoretical, intuitive sense of which actions count as sanctionable by considering the following pair of cases alongside the standard luck egalitarian evaluation of said cases.

> *Coyote Bite*: When she was a baby, Agent was bitten by a coyote while momentarily left unattended on the front stoop. As a result of the bite, she now suffers from untreatable pain. She otherwise lives a life identical in quality to that of her neighbor, Compensator.

In this case, Agent is left worse-off than Compensator by her injury. Additionally, there is no sanctionable choice on her part that would justify this inequality. She did not cause her own injury, and even if she did, she would not have been responsible for that injury given that she was a baby. And by hypothesis,

---

made by the worse-off party that justifies inequality, with the "sanction" in question being the party losing any claim to redistributive transfers. For example, if making avoidable gambles is what justifies inequality, then someone who invests in a risky stock and ends up worse-off as a result chooses sanctionably, even though she does not commit any sort of wrongdoing by investing.

Some might take this presentation of luck egalitarianism to be idiosyncratic, as it follows my favored characterization of the position (as developed in Spafford, "Luck Egalitarianism Without Moral Tyranny" and *Social Anarchism and the Rejection of Moral Tyranny*), where this characterization departs from more orthodox characterizations of luck egalitarianism. For example, an anonymous reviewer objects that luck egalitarians do not want to sanction/penalize choices; rather, they want to hold people responsible for their choices. However, the paper uses my formulation because it is a maximally general way of defining luck egalitarianism that is inclusive of the reviewer's formulation while also accommodating other articulations of the position. According to the reviewer's orthodox formulation of luck egalitarianism, an inequality is just iff (roughly) it resulted from the worse-off party's choice(s). By contrast, luck egalitarians like Jens Damgaard Thaysen and Andrea Albertsen (discussed below) take an inequality to be just if the worse-off party created a cost that must be borne by someone. See Thaysen and Albertsen, "When Bad Things Happen to Good People." The advantage of my proposed formulation of luck egalitarianism is that it can declare both of these rival views to be variants of luck egalitarianism, differing only with respect to the account of sanctionable choice they incorporate into the broader theory. On the orthodox formulation, a person chooses sanctionably when she leaves herself worse-off than another; on Thaysen and Albertsen's view, she chooses sanctionably when she creates a cost that someone has to internalize. But despite this point of disagreement, both proposals would qualify as variants of luck egalitarianism on my proposed account. Finally, one might also note that my proposed formulation resembles the general formulation of luck egalitarianism posited by G. A. Cohen (*Rescuing Justice and Equality*, 7), which similarly departs from the orthodox formulation posited by the reviewer. For various alternative statements of luck egalitarianism, see Temkin, *Inequality*, 13; Vallentyne, "Brute Luck and Responsibility," 58; Arneson, "Liberalism, Capitalism, and 'Socialist' Principles," 243; and Lippert-Rasmussen, *Luck Egalitarianism*, 1.

she has not done anything else that might play the appropriate justificatory role. Thus, luck egalitarians would hold that the inequality between Agent and Compensator is unjust, with justice thereby requiring that Compensator make transfers to Agent such that Agent is partially compensated for her suffering. Specifically, if there are n total people including Agent and Compensator, they are otherwise equally situated, and the loss of advantage that Agent incurs is equal to x, then justice would require that each person make a transfer to Agent such that all persons incur a cost equal to x/n, thereby realizing an equal distribution of advantage.[29] Thus, given the simplifying assumption that there are no other parties present in Coyote Bite besides those who are named—a simplifying assumption that will also be made for all other cases discussed in this paper—Compensator must make transfers up to the point where both she and Agent end up with $x/2$ fewer units of advantage than they would have had absent the coyote attack.

This assessment can be contrasted with that of the following case.

> *Coyote Wrestling*: Despite knowing the dangers involved, Agent decides to wrestle a coyote for fun. In the process, she is bitten by the coyote, and the resulting injury leaves her with untreatable pain. She otherwise lives a life that is identical in quality to that of her neighbor, Compensator.

In this case too, Agent is left worse-off than Compensator by her injury. However, unlike in Coyote Bite, luck egalitarians will say that Agent made a sanctionable choice when she decided to wrestle the coyote, where this choice renders the inequality between Agent and Compensator just. By declaring the inequality just, they are thereby able to avoid the implausible conclusion that justice requires Compensator making costly transfers to Agent. Given that it would seemingly be unfair if Agent and Compensator had to equally share the costs generated by Agent's choice, a theory of justice that prescribes such a distribution—e.g., one that demands strict equality—lacks extensional adequacy.

29 This assumes that the marginal advantage produced by the transfers is neither diminishing nor increasing. There is also a further assumption that what matters is equality across lifetimes, with the posited transfer making Agent better-off later in her life to make up for her being worse-off than others prior to the transfer. There is some debate among egalitarians over whether equality must be realized across entire lives (see Dworkin, *Sovereign Virtue*, 89; Lippert-Rasmussen, *Luck Egalitarianism*, 154–56; and Segall, *Why Inequality Matters*, 86–89) or across shorter segments of time as well (see McKerlie, "Equality and Time" and "Justice Between the Old and the Young"; and Temkin, *Inequality*). However, as Segall notes, those in the latter camp still hold that egalitarians are concerned with lifetime equality such that one has reason to realize it at the expense of creating inequalities across shorter spans of time (*Why Inequality Matters*, 84).

By contrast, luck egalitarianism preserves its extensional adequacy by requiring that Agent internalize these costs.

There are various accounts of sanctionable choice that yield the judgment that Agent chooses sanctionably in Coyote Wrestling but not Coyote Bite. For example, the classical account holds that a person chooses sanctionably iff her choice leaves her worse-off than others and she is responsible for this outcome.[30] However, in light of the various concerns that have been raised about this account, one might instead follow Jens Damgaard Thaysen and Andreas Albertsen, who contend that sanctionable choices are ones that generate costs that must be borne by someone.[31] More precisely, these choices generate a smaller quantity of total advantage than some alternative choice available to the agent. In other words, they are choices that make it such that someone must end up with less advantage than they otherwise could have had.

By declaring choices of this kind sanctionable, luck egalitarians ensure that it is the agent who ends up internalizing these imposed costs rather than anyone else. For example, in Coyote Wrestling, Agent chooses sanctionably because she makes it such that there is less total advantage to go around. Had she made a different choice, everyone could have lived rich, pain-free lives. However, by wrestling a coyote, she makes it so that someone must be left comparatively worse-off: either Agent's life goes worse due the resultant pain, or she is compensated for her suffering but at Compensator's expense. Given that someone must bear a cost as the result of her choice, that choice is sanctionable.

30 See Dworkin, *Sovereign Virtue*, 73; Cohen, *Why Not Socialism?* 17–18; and Lippert Rasmussen, *Luck Egalitarianism*, 2, 5.

31 Thaysen and Albertsen, "When Bad Things Happen to Good People." For an influential objection to the classical account, see the "boring problem" raised by Susan Hurley, *Justice, Luck, and Knowledge*, 160–61. Granted, I have recently argued that the boring problem is not, in fact, a problem for luck egalitarians (Spafford, review of *Strokes of Luck*, 432–33). However, there are other significant challenges to the classical account that are much harder to dismiss. See, e.g., Olsaretti, "Responsibility and the Consequences of Choice." Regarding the choice to adopt Thaysen's and Albertsen's account here, it should be noted that I actually defend a rival account of sanctionable choice in Spafford, "Luck Egalitarianism Without Moral Tyranny" and *Social Anarchism and the Rejection of Moral Tyranny*. That said, my favored account largely aligns with Thaysen's and Albertsen's while being a fair bit more complex; thus, for the sake of simplicity, I use their account in what follows (though the account is ultimately adjusted such that it more resembles mine for reasons discussed in section 6 below). (See also note 34 below.) Granted, there are other accounts one might adopt instead. For example, the boring problem has led Gerald Lang to reject the classical account of sanctionable choice and propose an original alternative ("How Interesting Is the 'Boring Problem' for Luck Egalitarianism?" and *Strokes of Luck*). However, given his doubts about the adequacy of his own proposal (*Strokes of Luck*, 196–99), one might reasonably endorse Thaysen's and Albertsen's account (either independently or as a comparatively-easy-to-work-with approximation of my own favored view).

Thus, on this interpretation of luck egalitarianism, any inequality generated by her internalizing the generated costs is just.

So far, this discussion has all attempted to describe the particular shares of advantage that luck egalitarians think that a person should distribute to others, *ceteris paribus*. In other words, it has discussed the allocative aspect of luck egalitarianism. However, to show that the position is a rights-based allocative theory, it must show that luck egalitarianism assigns rights on the basis of its prescribed shares. To see why this is a plausible interpretation of the theory, note that the characterization of the theory above already suggests that the theory is appropriately stated in terms of rights. Specifically, note that the apparent questions that luck egalitarians seek to answer are (a) whether Compensator acquires a *duty* to eliminate inequality in Coyote Bite/Coyote Wrestling by either transferring advantage to Agent or simply leveling down by reducing her own share of advantage and (b) whether Agent has a *duty* to refrain from acting in ways that would involuntarily transfer advantage from Compensator or level down by reducing Compensator's share of advantage. After all, the point of the theory is seemingly to determine the permissibility of redistribution, where this permissibility will be a function of persons' various permissions and duties to redistribute. Thus, luck egalitarianism must be construed as a theory of duties. Of course, duties do not necessarily entail the existence of some correlative right, as they might be nondirected, i.e., not owed to any particular person(s). However, it seems plausible to think that luck egalitarianism's posited duties *are* directed given the fact that it seems that Compensator would *wrong* Agent in Coyote Bite if she were to refuse to transfer any advantage to Agent. To fail to discharge this duty would be to deny Agent her just share and also benefit at Agent's expense. Such a choice would seemingly give Agent a basis for complaint against Compensator and render feelings of resentment on her part apt—where both of these facts are signature features of Agent having a right infringed. And for identical reasons, it seems that Agent would wrong Compensator in Coyote Wrestling if she were to try to restore equality by either taking some of Compensator's holdings or otherwise diminishing Compensator's advantage.[32]

32 Note that pure telic luck egalitarians would likely not endorse a rights-based interpretation of luck egalitarianism. For such luck egalitarians, inequality in the absence of sanctionable choice is a bad-making property of states of affairs that consequentialist agents must weigh against their various good-making properties when determining which realizable state of affairs has the maximal quantity of moral value. Given this view, telic egalitarians tend not to posit that persons have any sort of right to equality (for the original introduction of this distinction, see Parfit, "Equality or Priority?" 84). However, their views might end up being coextensive with the proposed theory depending on how much comparative moral weight they assign to the badness of inequality versus good-making properties like the

While luck egalitarianism is not typically articulated using the language of rights, I have recently defended a luck egalitarian theory that rejects the existence of property claims and instead assigns persons natural luck egalitarian rights over resources. On my view,

> Each person [is assigned] a set of claims such that the luck egalitarian principle would be satisfied if all persons respected the claims of others—that is, any inequality would appropriately correspond to some sanctionable choice on the part of the worse-off individuals. Or, to put this point slightly differently, each person would have a claim against anyone else using an unowned resource in some way if and only if that use would leave her with less than her appropriate share of advantage, where her *appropriate* share is either (a) equal to the respective shares of those who have not yet chosen sanctionably if she has also not yet chosen sanctionably or (b) adjusted downward from this value if she has chosen sanctionably.[33]

Notably, condition b implies that sanctionable choice leads to rights forfeiture. For example, prior to wrestling the coyote, Agent has a set of claims against others using resources in ways that will leave her worse-off than them. However, after her choice, she loses some subset of these claims. Specifically, if one endorses the general approach to sanctionable choice posited above, she loses a subset of claims such that justice requires her fully internalizing the costs that she produces (such that Compensator does not have to internalize those costs).

When luck egalitarianism is presented in these terms—namely, as a theory of egalitarian rights that are forfeited when agents choose sanctionably by generating costs—then it becomes apparent that it is a theory of the same kind as the compensation thesis. Specifically, both theories assign rights and duties such that everyone discharging their respective duties will realize a desired

---

total quantity of advantage that persons possess. Note that even if persons have directed luck egalitarian duties to realize particular distributive states of affairs, these are just *pro tanto* considerations that might be overridden by other consequentialist considerations such as the effect that discharging those duties would have on the total quantity of advantage. Typically, this overriding occurs when the goodness of the consequentialist considerations is x times greater than the badness of infringing the right. Thus, very roughly, if telic egalitarians declare the badness of luck-based inequality to be x times weightier than the goodness of rival considerations, their judgments of all-things-considered permissibility will be coextensive with those of rights-based luck egalitarians. That said, telic luck egalitarianism would fail to capture the apparent directedness of redistributive duties just described in the main text, making rights-based luck egalitarianism the superior theory even if the two views prescribe/proscribe the same actions.

33 Spafford, *Social Anarchism and the Rejection of Moral Tyranny*, 52.

distribution of advantage across at least certain persons. The question then becomes whether these two theories are compatible. The subsequent section will answer this question in the negative: the two theories imply incompatible judgments when applied to cases of blameless wronging like Consent Evidence and Forfeiture Evidence.

6. The Incompatibility of Luck Egalitarianism and the Compensation Thesis

According to the interpretation of luck egalitarianism provided just above, an agent loses claims against others using resources in advantage-diminishing ways when the agent makes a choice that generates costs. However, the proposal is arguably imprecise, as a natural amendment to this statement is that the agent loses these claims only if the evidence she possesses (and the evidence that she reasonably ought to have gathered) suggests that her action will generate such costs. To motivate this thought, consider the following case.

> *Coyote Surprise*: Agent decides to go for a run in an urban area known for its safety. However, in an unprecedented event, a coyote that had just escaped from a nearby zoo leaps out of a bush and bites Agent, and the resulting injury leaves her with untreatable pain. She otherwise lives a life that is identical in quality to her neighbor, Compensator.

Intuitively, it seems that Agent's normative status in Coyote Surprise is of a kind with Agent's status in Coyote Bite rather than with Agent's status in Coyote Wrestling. Specifically, in both Coyote Bite and Coyote Surprise, it seems unfair if Agent is left worse-off than Compensator. By contrast, it does not seem unfair if Agent is left worse-off than Compensator in Coyote Wrestling—and in fact, it would seemingly be unfair if Compensator had to compensate Agent for her suffering.

If this intuitive judgment is determinative, then one must adjust the proposed account of sanctionable choice to accommodate it. As currently stated, the account implies that Agent chooses sanctionably in Coyote Surprise, as she acts in a way that generates costs: had she not chosen to go for a run, the coyote, by hypothesis, would not have bitten anyone, with no distributable costs thereby being generated. To avoid this implication, the account must be amended such that Agent's choice in Coyote Surprise does not count as sanctionable. An extensionally adequate account will thus deliver this judgment while still preserving the implications that (a) Agent chooses sanctionably in Coyote Wrestling and (b) Agent does not choose sanctionably in Coyote Bite.

To deliver these results, one can adjust the account by making sanctionable choice a function of *blameworthy* cost imposition, with evidence-based foreseeability being a necessary condition of blameworthiness.[34] On this approach, the explanation for why it is fair to hold Agent accountable for the imposed costs in Coyote Wrestling is that Agent is to blame for those costs (as she both causes them, and the possessed and relevant evidence suggests that her choice will bring them about). Similarly, it would be unfair to hold Agent accountable in Coyote Bite because she is not to blame for those costs (since she did not cause them). And while Agent *does* cause the costs in Coyote Surprise, that imposition of costs is also not blameworthy (due to the fact that given both her

34 This proposal has been endorsed by various luck egalitarians such as Carl Knight ("Egalitarian Justice and Expected Value" and "An Argument for All-Luck Egalitarianism"), and it is one that I defend elsewhere (Spafford, "Luck Egalitarianism Without Moral Tyranny" and *Social Anarchism and the Rejection of Moral Tyranny*, ch. 6). More precisely, I make sanctionable choice depend upon whether the agent has *reason to expect that* her action will generate costs (where the expected value of an action might be positive even if it is foreseeable that it might impose costs). However, making mere foreseeability a necessary condition of sanctionable choice is a weaker assumption than adopting my full account, so this argument incorporates the former position rather than the latter.

The posited proviso could be made weaker still by making foreseeability alone a necessary condition of sanctionable choice (rather than blameworthiness more generally) without compromising the validity of the paper's argument. However, the theoretical advantage of the more general proviso is that it seems less ad hoc, as it provides a unifying explanation of *why* foreseeability is a necessary condition of sanctionable choice that is intuitively plausible (as it is plausible that one must be blameworthy if one is to be left comparatively worse-off). The disadvantage of positing the more general proviso is that it will not be applicable to various other accounts of sanctionable choice. Note that a blameworthiness proviso seems apt when appended to the account of sanctionable choice endorsed by the paper (namely, Thaysen's and Albertsen's account); however, it would not seem apt if one adopted instead an account that declared a choice sanctionable iff it resulted in the chooser ending up worse-off than another party. This is because imposing avoidable costs that must now be distributed seems like a form of wrongdoing that might therefore be blameworthy while simply leaving oneself worse-off is not obviously wrongdoing and thus is not the sort of choice that can be blameworthy. That said, rival accounts of sanctionable choice will still seemingly need to posit some analog to the blameworthiness proviso to avoid objectionable implications (e.g., that an agent chooses sanctionably if she leaves herself worse-off in a way that she could not have foreseen). For example, while Dworkin articulates his version of luck egalitarianism very differently from how things are presented here, he would hold that (a) an inequality between Agent and Compensator in Coyote Surprise—i.e., holding Agent responsible in this case—would be justified only if Agent had declined an opportunity to insure against this outcome and (b) Agent has such an opportunity to insure only if she knows the risk of the outcome occurring (*Sovereign Virtue*, 77). Given that one cannot know the risk of an unforeseeable outcome occurring, it follows that foreseeability is a necessary condition of holding individuals responsible on Dworkin's view, where this is akin to the posited blameworthiness proviso.

evidence and the evidence more broadly available to her, she could not foresee that her action would impose those costs).

If this is correct, then the rights-based approach to luck egalitarianism can be understood as asserting the following: in any case where some person carries out some cost-imposing action ϕ, her *rights to advantage*—i.e., her permissions to act in ways that would yield some specified share of advantage and her claims against actions that would reduce this share of advantage—are uniquely lost only if either her possessed evidence or the relevant evidence suggests that ϕ-ing will generate costs that must ultimately be distributed across persons.[35] Thus, Agent would not uniquely lose any rights to advantage in Coyote Surprise but would uniquely lose such rights in Coyote Wrestling.

Contrast this result with that implied by the compensation thesis. The latter holds that when Aggressor ϕs in cases like Consent Evidence and Forfeiture Evidence, she uniquely loses those rights to advantage that would cause her to fully internalize the costs that she generates. Further, Aggressor loses these rights even though neither her evidence nor the relevant evidence suggests that ϕ-ing will impose these costs. However, note that this conclusion directly contradicts the luck egalitarian thesis presented in the previous paragraph, namely, that blameworthiness is a necessary condition of a person uniquely losing rights to advantage. Thus, accepting the compensation thesis requires rejecting rights-based luck egalitarianism.[36] This is a high theoretical cost for a rights theorist to incur, as it yokes the adequacy of the thesis to the falsity of a plausible interpretation of a popular theory of distributive justice. While this might not trouble libertarians and others who explicitly reject luck egalitarianism,

35 The reason for the qualifier 'unique' is that, strictly speaking, luck egalitarians will hold that a person's rights set *is* diminished in any case where she either imposes or incurs costs, e.g., Coyote Surprise and Coyote Bite. However, it is diminished along with everyone else's set of rights, as each person's set of rights are adjusted to ensure that the imposed/incurred costs are evenly distributed (i.e., full compliance with everyone's adjusted rights will result in each person receiving a share of advantage that is smaller than the one she would have otherwise received and smaller to an equal degree).

36 One might try to avoid contradiction by amending luck egalitarianism's necessary condition of unique rights loss such that is disjunctive: a person loses rights to advantage in virtue of ϕ-ing when *either* the relevant evidence indicates that ϕ-ing will impose costs *or* ϕ-ing infringes someone's rights. However, this proposal seems unacceptably arbitrary. Why is it that in some cases (e.g., Coyote Surprise), being blameless for ϕ-ing precludes Agent from forfeiting rights to advantage that she otherwise would have forfeited, but in other cases (e.g., Consent Evidence), blamelessness does not preclude Aggressor from forfeiting such rights? Absent a principled explanation of this supposed difference, one cannot plausibly maintain that blamelessness sometimes precludes forfeiture but does not have this effect in cases of rights infringement.

rights theorists without such commitments might be reluctant to take on such a significant commitment vis-à-vis distributive justice.

More directly, to affirm the compensation thesis is to deny that blamelessness insulates agents from uniquely losing rights to advantage. Thus, proponents of the solution cannot appeal to Agent's blamelessness in Coyote Surprise to support the exculpatory judgment that Agent does not have to fully internalize the costs she incurs as a result of going for a run. Absent this theoretical resource, they may find themselves committed to the proposition that justice requires Agent fully internalizing the costs of being attacked by a coyote in Coyote Surprise—a result that many might find troubling even setting aside the more general question of whether the compensation thesis is incompatible with rights-based luck egalitarianism.

7. AGAINST THE COMPARATIVE FAIRNESS SOLUTION

Having introduced luck egalitarianism, it is now possible to explain why the comparative fairness solution does not adequately resolve the trilemma of section 1. Recall that this solution holds that it is in fact fair for agents to internalize the costs of blameless ϕ-ing. Granted, this position is *prima facie* implausible, as even Jorgensen concedes that it seems unfair to make agents like Aggressor bear the costs in cases of blameless rights infringement. However, as noted above, Jorgensen suggests that it is ultimately fair to make Aggressor internalize the costs of blameless ϕ-ing because this is the least unfair option available (since the alternative of making Claimholder internalize those costs is even less fair).

This section argues that this inference rests on a false premise: if one accepts the luck egalitarian position, then one can—and indeed must—endorse a third, alternative prescription vis-à-vis Consent Evidence and Forfeiture Evidence. Further, this alternative appears to be a fairer option than either of the ones that Jorgensen considers. Thus, one cannot defend the comparative fairness solution by appealing to the comparative unacceptability of the available alternative(s).

To see what luck egalitarianism implies vis-à-vis Consent Evidence/Forfeiture Evidence, recall first what it prescribes in cases like Coyote Bite and Coyote Surprise, wherein an agent imposes costs but is not blameworthy for doing so. As noted in section 5, if all parties start out equally situated, then any imposed costs must be equally distributed across persons. Thus, in both cases, if the pain of the coyote bite reduces Agent's advantage by x units, then Compensator must make transfers to Agent such that they each incur a cost of $x/2$ (since there are only two parties present in the scenario). This distribution of costs ensures that the overall distribution of advantage remains equal, thereby

avoiding the injustice of an inequality that cannot be vindicated by some sanctionable choice on the part of the worse-off.

Given that Consent Evidence and Forfeiture Evidence are of a kind with Coyote Bite and Coyote Surprise—i.e., they are cases in which an imposed cost that does not correspond to some sanctionable choice must be distributed—luck egalitarians would prescribe the same distribution of costs: Aggressor and Claimholder should each internalize half of the total costs of Aggressor ϕ-ing.[37] Thus, luck egalitarianism reveals (via its prescription) that there is an alternative way of distributing the costs of blameless ϕ-ing beyond having either Aggressor or Claimholder fully internalize them. Further, note that this proposed distribution seems fairer than either of the proposed alternatives. Given that no one is to blame for the imposed costs, it would be arbitrary—and thus unfair—to make only Aggressor or only Claimholder incur those costs. By contrast, an equal division of costs across persons seemingly avoids this arbitrariness/unfairness.

Finally, note that this proposal also resolves Jorgensen's worry that it would be comparatively less fair for Claimholder to internalize the costs of Aggressor ϕ-ing in cases where Claimholder, for social reasons, is prone to having her rights infringed. Recall the case from section 3 wherein Claimholder ends up having her claims infringed by both Aggressor blamelessly ϕ-ing and, later, Infringer blamelessly ϕ-ing. The plausible suggestion above was that it would be fairer for Aggressor and Infringer to internalize their respective imposed costs than have Claimholder internalize both sets of costs. However, again, to consider only these two options is to propose a false dilemma, as the luck egalitarian would endorse a third option that seems fairer than either of the aforementioned ones: for each instance of blameless rights infringement, each party should internalize an equal share of the costs. In other words, Aggressor, Infringer, and Claimholder would each absorb one-third of the costs of Aggressor ϕ-ing and one-third of the costs of Infringer ϕ-ing. This outcome precludes Claimholder from having to bear any special burden resulting from others' propensity to infringe her rights.

Given that Aggressor internalizing the costs of blameless ϕ-ing is not the least unfair option available, Jorgensen cannot infer that it is fair *tout court*. Thus, Jorgensen's argument for the comparative fairness solution must be rejected. Further, one should reject the solution itself, as it is *prima facie* implausible and

37 For a related suggestion, see Preda, who proposes that in certain cases of permissible infringement where someone infringes a claim as an unavoidable side effect of enforcing a different claim to some state of affairs obtaining, the costs of that infringement should be shared by everyone who had a correlative duty to realize that state of affairs ("Are There Any Conflicts of Rights?" 686–87).

is no longer backed by a supporting argument capable of overriding that *prima facie* judgment.[38] This rejection, when paired with section 2's rejection of the noninfringement solution, leaves only one remaining option for resolving the trilemma of section 1: reject the compensation thesis.

8. ARGUMENTS FOR THE COMPENSATION THESIS

The previous sections have presented two distinct arguments against the compensation thesis. First, there was the incompatibilist argument of sections 4–6, which tried to show that those who endorse the thesis must pay the high theoretical price of rejecting (rights-based) luck egalitarianism. Second, there was the argument by elimination: to resolve the trilemma of section 1, one must choose between endorsing the noninfringement solution, endorsing the comparative fairness solution, or rejecting the compensation thesis, where only the last-mentioned option is theoretically acceptable. However, these arguments must be weighed against the positive arguments that can be marshalled in support of the compensation thesis. This section considers six such arguments and argues that none of them succeed.

To begin, one might argue for the compensation thesis—and simultaneously object to the incompatibility argument of section 6—by suggesting that, actually, luck egalitarianism *implies* the judgment that Aggressor must uniquely compensate Claimholder in cases like Consent Evidence and Forfeiture Evidence, which is to say that luck egalitarianism implies the compensation thesis. This argument takes the signature commitment of luck egalitarianism to be holding people responsible for their choices by making them internalize the resulting costs. For example, when agents gamble and lose, luck egalitarians classically hold that any resulting inequality is just, with the losers thereby not being entitled to any compensation for their bad *option luck*.[39] Further, to compensate them would be *unjust*, as that would make others absorb the costs of their reckless choices.

38 Jorgensen also gives a second quick argument in defense of the comparative fairness solution. Specifically, she notes that it would be appropriate for Claimholder to internalize the costs of Aggressor ϕ-ing if Claimholder had consented to that ϕ-ing; however, given that Claimholder has not consented in cases of blameless wronging, it would be inappropriate for her to internalize those costs—where this leaves only Aggressor to absorb the costs (Jorgensen Bolinger, "Moral Risk and Communicating Consent," 202). The problem with this argument is that just because Claimholder consenting is a sufficient condition of it being appropriate for Claimholder to internalize the costs of ϕ-ing, it does not follow that it is also a necessary condition.

39 Where option luck contrasts with *brute luck*, i.e., inequality-grounding costs that are the result of luck but not avoidable gambles (e.g., the costs imposed by congenital health conditions).

However, when applied to Consent Evidence and Forfeiture Evidence, luck egalitarianism would then imply that Aggressor should absorb the costs of her choice to ϕ, as this outcome would be just (because it is simply making Aggressor internalize the costs resulting from the gamble she made by ϕ-ing), with the alternative being unjust (as Claimholder would then be forced to absorb the costs created by Aggressor).[40] Further, the way to make it such that Aggressor internalizes the costs of her wrongdoing is by having her pay compensation to Claimholder. Thus, contrary to the foregoing argument, luck egalitarianism actually supports the compensation thesis's contention that wrongdoers have a unique duty to compensate the wronged parties for any harm inflicted.[41]

The reply to this argument begins with the observation that there are many rival interpretations of luck egalitarianism, each with its own distinct set of implications. Some of these interpretations offer rival accounts of what qualifies as advantage (i.e., what it is that must be distributed equally in the absence of sanctionable choice); others differ with respect to whether they declare an equality unjust when one party has chosen sanctionably (as discussed in note 28 above). For these purposes, the relevant point of interpretive disagreement is over which choices qualify as sanctionable. Specifically, note that the just presented luck egalitarian argument for the compensation thesis presupposes the

40 Note that this judgment also aligns with the prescriptions of the *responsibility account*, a widely endorsed position in the literature on self-defense. According to those who defend some version of this account (e.g., McMahan, "The Basis of Moral Liability to Defensive Killing"; Otsuka, "Killing the Innocent in Self-Defense," 91; and Gordon-Solmon, "What Makes a Person Liable to Defensive Harm?"), when *P* threatens *Q* with wrongful lethal harm, fairness requires that *P* internalize that harm rather than *Q*. Thus, *Q* can permissibly defend herself by killing *P* in such circumstances. This seems to be an implication of the more general luck egalitarian view that fairness requires agents internalizing the costs they create (as noted by Gordon-Solmon, "What Makes a Person Liable to Defensive Harm?" 546). If this is right, the advocate of the proposed argument/objection might hold that the compensation thesis and the responsibility account are both facets of luck egalitarianism. Notably, similar remarks might apply to Jorgensen's fairness-based account of self-defense, which builds on the responsibility account by positing that *Q* can permissibly defend herself if *P either* threatens *Q* with wrongful harm *or* gives *Q* evidence that she has threatened *Q* with harm (along with various constraints that cannot be listed here). See Jorgensen Bolinger, "The Moral Grounds of Reasonably Mistaken Self-Defense," 147. While this account diverges from standard responsibility accounts in that it makes permissible self-defense a function of evidence versus mere harm, this extension might be seen as aligning with evidentialist versions of luck egalitarianism such as the one that I endorse ("Luck Egalitarianism without Moral Tyranny" and *Social Anarchism and the Rejection of Moral Tyranny*). If both Jorgensen's responsibility account and her defense of the compensation thesis do indeed follow from a more general luck egalitarian approach to justice, that would make her views of consent and self-defense coherent in a way that is theoretically attractive.

41 I am grateful to an anonymous reviewer for raising this argument/objection.

following interpretation of sanctionable choice: an agent chooses sanctionably if she makes it such that an otherwise avoidable cost must now be absorbed by some person(s). Thus, when Aggressor generates a distributable cost in Consent Evidence/Forfeiture Evidence, she is held to have made a sanctionable choice, making the just outcome one where she is left worse-off than Claimholder.

This proposed account of sanctionable choice is seemingly identical to the one proposed in section 5. There, following Thaysen and Albertsen, it was suggested that a sanctionable choice is one that generates costs that must be borne by someone—with this account supporting the just presented argument's conclusion that Aggressor chooses sanctionably in Consent Evidence/Forfeiture Evidence. However, recall that section 6 argued that an interpretation of luck egalitarianism that incorporates this account of sanctionable choice is extensionally inadequate, as it generates unacceptable results in Coyote Surprise. In that case, it was posited that Agent does not choose sanctionably (i.e., it would be unjust if she were left worse-off than others) despite the fact that she made a choice that generated costs. Given that the posited interpretation of luck egalitarianism/sanctionable choice yields an unacceptable result in Coyote Surprise, it was argued that one must reinterpret the notion of sanctionable choice such that sanctionable choices include only those where the agent is *blameworthy* for producing some cost (where blameworthiness is a function of what is foreseeable given the available evidence). Finally, since Aggressor is not blameworthy in either Consent Evidence or Forfeiture Evidence, this extensional adequacy-improving (re)interpretation of luck egalitarianism implies that she does not choose sanctionably and thus that it would be unjust if she had to fully internalize the cost of ϕ-ing. In short, while some interpretations of luck egalitarianism do imply the compensation thesis—and thus function as arguments for the thesis—these interpretations are unacceptable due to being extensionally inadequate.[42] The first proposed argument for the compensation

42 What, then, is the relationship between luck egalitarianism and the responsibility account discussed in note 40 above? Answering this question is complicated by the fact that proponents of the responsibility account are typically interested in determining how to distribute only *indivisible* costs, whereas the proposed interpretation of luck egalitarianism generally presupposes that costs are divisible, e.g., when it prescribes that costs are to be equally shared in cases of blameless cost imposition. One possibility is that the proposed interpretation of luck egalitarianism is incompatible with the responsibility account, as the former might imply that if *P* threatens *Q* with wrongful indivisible harm—but is not blameworthy for doing so, i.e., her choice is not a sanctionable one—then a coin should be flipped to determine who should incur the cost. Such incompatibility would mean that the argument from Coyote Surprise would have to be weighed against whatever reasons there are for accepting the responsibility account. Alternatively, one might take luck egalitarianism and the responsibility account to have distinct subject matters, with the former applying only to cases where costs are divisible while the latter applies only to

thesis is therefore unsound (and also fails as an objection to the incompatibility argument of section 6).

Second, one might endorse the compensation thesis on the more general grounds that it would be unfair if the rights holder had to incur costs generated by someone violating her rights.[43] However, the discussion of the previous section serves to undermine this proposal. While it might initially seem unfair to impose costs on Claimholder in cases like Consent Evidence and Forfeiture Evidence, this judgment is substantially softened once one considers that the only alternative is making a blameless Aggressor fully internalize the costs—an alternative that runs contrary to both luck egalitarianism and the broad set of intuitive judgments to which its proponents appeal. As discussed in section 7, once one notes that it seems equally unfair for either a blameless Aggressor or a wronged Claimholder to fully internalize the costs of Aggressor's blameless action, then having Claimholder internalize half of the imposed costs does not seem unfair.

Third, one might appeal to Loren Lomasky's argument for the compensation thesis.[44] Unfortunately, this argument rests partially on Lomasky's theory about the grounds of rights, and fully recapitulating and critically engaging with this theory would go beyond the scope of this paper. However, a more general version of the argument can be posited that does not presuppose Lomasky's grounding theory. On this reconstruction, when assessing who should internalize the costs of ϕ-ing, one must look to whose projects are advanced by the ϕ-ing (where a project is roughly an end that a person is committed to realizing via some long-term, self-conception-modifying, life-structuring plan).[45] For example, in the Permissible Infringement case, Lomasky notes that it is only Hiker whose ends are advanced by the use of Owner's cabin. Again, setting aside the exact details of Lomasky's account, he suggests that it would be inappropriate for Owner to then incur the costs associated with Hiker using the cabin to survive for the sake of pursuing Hiker's various projects. Indeed, one might generalize this judgment by positing that it is unfair to make a victim of rights infringement internalize the costs of said infringement when the benefits

cases where costs are indivisible (and those costs will be imposed via one party wronging another). This approach would render the two positions compatible, as luck egalitarianism would imply that costs should be shared equally in Forfeiture Evidence without the further implication that it would be wrongful for Claimholder to kill Aggressor if Aggressor would otherwise kill her (i.e., if the costs to be distributed were indivisible). Unfortunately, a comparative assessment of these two possibilities cannot be provided here.

43 For an example of this worry, see Davis, "Rights, Permission, and Compensation," 381.

44 Lomasky, "Compensation and the Bounds of Rights."

45 Lomasky, *Rights Angles*, 50.

of that infringement are internalized by the perpetrator of the infringement.[46] Given that such a distribution of costs would seemingly be unfair even if the cost-imposing action were not a rights infringement, this Lomasky-inspired principle seems plausible.

Despite its plausibility, this principle fails to support the compensation thesis for two related reasons. First, there will be many cases of infringement where the beneficiary of the action is someone other than the infringing agent. For example, suppose that in Consent Evidence, Aggressor ϕs with the intention of benefitting both Claimholder and a third party, Beneficiary, the latter of whom actually benefits (unlike Claimholder, who incurs a cost as per the initial description of the case). Given that Aggressor does not intend to benefit from ϕ-ing and does not internalize any benefits, the Lomasky-inspired principle would not support the judgment that Aggressor should internalize the costs of ϕ-ing. Rather, seemingly either Beneficiary or both Claimholder and Beneficiary should internalize the costs given that they are the (intended) beneficiaries of the action.[47]

The second problem with the argument from Lomasky's principle is that it fails to take into consideration how luck egalitarianism bears upon the question of whose projects are advanced by a given act of ϕ-ing. Note that if luck egalitarianism is correct, the benefits of any given action are effectively socialized, as justice requires that any generated advantage be distributed in a way that realizes an equal distribution (or, more precisely, a distribution that is either equal or unequal in a way that corresponds to previous sanctionable choices). For example, suppose that Augmenter blamelessly wrongs Costbearer by ϕ-ing, where ϕ-ing produces y units of total distributable advantage but also imposes a cost of x on Costbearer. Further, suppose that (a) y is greater than x, (b) Augmenter and Costbearer are the only existing people (an assumption of all the cases in this paper), (c) they possess equal quantities of advantage, and (d) neither has previously made a sanctionable choice. How then should x and y be distributed? If the compensation thesis were false, the luck egalitarian would say that Costbearer's loss of x should first be offset using the gain of y (to restore equality), with the remainder of the gains from y being divided equally between Augmenter and Costbearer—i.e., both Augmenter and Costbearer receive $(y - x)/2$ units of advantage. This distribution of the costs and benefits of ϕ-ing would result in Augmenter and Costbearer ending up with equal quantities of advantage, which is what justice requires.

46 Jorgensen also seems to endorse something like this principle. See Jorgensen Bolinger, "Moral Risk and Communicating Consent," 203.

47 A similar argument is advanced by Montague, "Davis and Westen on Rights and Compensation," 393–94.

Suppose now that the compensation thesis is correct in addition to luck egalitarianism. Seemingly, one would first prescribe that advantage should be equalized in the way just described, as this is what luck egalitarianism requires. One must then factor in the compensation thesis, where its contribution must seemingly be that Augmenter must pay additional compensation to Costbearer even after the costs incurred by Costbearer have been offset by egalitarian redistribution. For example, if the thesis is taken to hold that Augmenter must pay full compensation to Costbearer, then one might hold that Augmenter must now make additional transfers to Costbearer such that Costbearer gains an additional x units of advantage.

There are two things to note about this proposal. First, this case functions as an additional reply to the fairness argument for the compensation thesis. Given that luck egalitarianism already ensures that Costbearer is fully compensated for her incurred costs—at least in cases where Augmenter's actions generate a positive quantity of distributable total advantage on net—then fairness does not seem to require her receiving any additional compensation from Augmenter. Second, the luck egalitarian prescription makes it unclear that Augmenter is the beneficiary of her ϕ-ing, thereby undermining the Lomasky-inspired rationale for endorsing the compensation thesis. Given that Augmenter must redistribute her gains in an equality-realizing way, it no longer seems that Augmenter's ϕ-ing is advancing her project; rather, it is advancing a collective egalitarian project, only the particular shape of which is influenced by Augmenter's choice. Granted, Augmenter might be the *de facto* beneficiary of her ϕ-ing if she declines to carry out the mandated luck egalitarian transfers to Costbearer. And in such a case, she would be obliged to make transfers to Costbearer. However, this obligation would be an implication of luck egalitarianism rather than one grounded in the appropriate distribution of the benefits of a rights infringement. Thus, the Lomasky-inspired argument does not seem to support the compensation thesis.

A fourth argument for the compensation thesis is that it is needed to adequately recognize Claimholder's status of having a claim against Aggressor ϕ-ing in Consent Evidence/Forfeiture Evidence. Note that a luck egalitarian theory of rights that excludes the compensation thesis yields identical distributive prescriptions in Consent Evidence, Forfeiture Evidence, and Coyote Surprise: in all three cases, the imposed costs of blameless choice (i.e., Aggressor's and Agent's choices) are distributed equally across all parties. Thus, the fact that Claimholder has a claim violated in Consent Evidence/Forfeiture Evidence—while Compensator does not suffer such violation in Coyote Surprise—seems to make no normative difference as far as the proposed compensation thesis-rejecting theory is concerned. However, it seems as though Claimholder's claim *should* make some normative difference.

This argument can be sidestepped by maintaining that even though Aggressor does not have to compensate Claimholder when Aggressor blamelessly wrongs her, Aggressor still has a duty *to apologize* to Claimholder for violating her claim. Such a remedial duty can then serve to distinguish the normative status of Compensator in Coyote Surprise from Claimholder in Consent Evidence/Forfeiture Evidence. While both agents impose costs, only Aggressor must apologize for her action, where this fact reflects Aggressor's unique species of wrongdoing. Thus, no duty of compensation is needed to distinguish agents who impose costs by wronging others (e.g., Aggressor in Consent Evidence) from those who simply impose costs (e.g., Agent in Coyote Surprise).

Fifth, one might be attracted to the compensation thesis because it serves other important normative functions. Consider, for example, an argument that Jorgensen makes in the context of defending the fairness of person Acculturated internalizing the full costs of ϕ-ing when (a) she (blamelessly) wrongs person Communicator because (b) Acculturated relies on a social convention that is morally problematic in some way (e.g., Communicator says, "You may not ϕ," but there is a conventional understanding in her society that this really means "You may ϕ" when uttered by women). Jorgensen argues that holding people responsible for acting on the evidence furnished by bad conventions—i.e., making them internalize the costs of their resulting wrongful actions—helps maintain and improve their responsiveness to moral reasons.[48] Jorgensen argues that when Acculturated relies on a bad convention, she fails to respond to the various moral reasons for not acting on that convention. This mistake is then corrected by holding Acculturated responsible as, by making Acculturated internalize the costs of her ϕ-ing, Acculturated and others are put "in a position to recognize that the [conventional] signals are bad ones."[49] Assuming that an adequate moral theory will be one that countenances this

48 Jorgensen Bolinger, "Moral Risk and Communicating Consent," 203. Here she cites Victoria McGreer and Philip Pettit, who argue that it is appropriate to hold agents responsible (via both blame and the associated reactive attitudes like resentment and indignation) in social contexts where third parties can make agents more disposed to respond to moral reasons via exhortation and expectation (McGreet and Pettit, "The Hard Problem of Responsibility," 177–79). However, as an interpretive matter, McGreer and Pettit seem more concerned with whether blame and emotions like resentment are *appropriate* than whether they are *useful* for cultivating reason-responsiveness. On their view, *the mere ascription of reason-responsiveness to agents* is a form of exhortation that cultivates the agents' capacity to respond to reasons, with blame/resentment merely being apt responses when agents fail to appropriately attend to those reasons. By contrast, Jorgensen seems to take blame to function as the exhortation that cultivates agents' moral capacities. This paper will focus on Jorgensen's construal, as it more directly supports endorsing the compensation thesis.

49 Jorgensen Bolinger, "Moral Risk and Communicating Consent," 203.

educative process, it follows that one should allow for Acculturated being held responsible for the costs of ϕ-ing.

Jorgensen's discussion focuses on cases where (a) Acculturated's available evidence suggests that ϕ-ing is permissible and (b) the fact that the evidence suggests this is morally problematic. However, her reasoning might equally apply to cases like Consent Evidence and Forfeiture Evidence, where the evidence is not morally problematic.[50] In cases of morally problematic conventions, the badness of those conventions gives agents reason to refrain from acting on the evidence furnished by those conventions. Holding them responsible for any resultant wrongs is then a way of drawing their attention to those overlooked reasons. Note, though, that the same reasoning might apply in Consent Evidence/Forfeiture Evidence: Aggressor has reason to refrain from acting on the evidence provided to her because it is not reliable. Thus, Jorgensen might maintain that Aggressor should be held responsible for ϕ-ing as a way of attuning her to this overlooked reason—where the compensation thesis must be affirmed if Aggressor is to be held responsible in this way.

The first thing to note about the foregoing argument is that it has an empirical element that might be questioned. Does discharging compensatory duties really educate wrongdoers in the posited way? Without collecting actual relevant data, it is hard to affirm this with certainty. Further, insofar as pre-empirical hypothesizing goes, it seems equally plausible that discharging a duty to apologize would educate at least as effectively as discharging a compensatory duty. Note that Jorgensen's hypothesis is that discharging compensatory duties will draw the agent's attention to the relevant reasons on which she should have acted. However, Aggressor apologizing to Claimholder would seemingly equally emphasize to Aggressor that she should not have ϕ-ed in cases like Consent Evidence and Forfeiture Evidence. In fact, the semantic nature of apology makes a duty of apology seem *better* suited for educating Aggressor than a duty of compensation. In theory, Aggressor could discharge a duty to compensate Claimholder without any understanding of what she did wrong or why she must carry out the act of compensation. By contrast, an adequate apology *would* require such an understanding. Thus, the educative effect of complying with a theory of rights that posits a duty of apology seems likely to be even greater than the effect of complying with the compensation thesis—though,

50 It is not fully clear whether this implication is intended by Jorgensen. Her proposal is specific to cases where morally problematic social conventions obtain. However, it follows immediately on the heels of a general defense of Aggressor having to internalize the costs of ϕ-ing in cases like Consent Evidence, suggesting that the proposal applies in these cases as well.

ultimately, empirical observation and experimentation would be needed to conclusively demonstrate this point.[51]

Finally, it is worth considering an argument for the compensation thesis developed in the literature on tort theory and advanced in various forms by Ernest Weinrib, Arthur Ripstein, and John Gardner.[52] In particular, the discussion will focus on Gardner's version of this argument, as he develops the point in greater detail and in explicitly moral terms. Consider the case where *A* owes *C* a duty not to ϕ. As Gardner presents the argument, there is some more fundamental moral reason *R* that grounds—and thus explains—*A*'s duty (60).[53] Now suppose that *A* breaches her duty by ϕ-ing. Gardner argues that in such a case, *A* no longer has a duty not to ϕ (as, presumably, agents cannot have duties to do what cannot be done) (59). However, while the duty is negated, Gardner argues that *R* persists. First, it persists in the sense that *A* still has reason not to ϕ *were she able to do so*, e.g., if she had the power to change the past (63). More importantly, it persists in the sense that the same reason for refraining from ϕ-ing can also be a reason to undertake other actions now that *A* has ϕ-ed. Specifically, it will (often) be a reason to compensate *C* via some compensatory action ψ (60–61), as compensating at least partially conforms to *R* now that full conformance is impossible. For example, Gardner suggests that everyone who rides a bus might have a duty to pay the fare up front because they all share the following reason to pay: paying "helps to see to it that the bus company gets paid for the services it provides, and hence is . . . encouraged to provide them" (60). If someone then forgets to pay the fare, she no longer has a duty to pay the fare before riding, but the posited grounds of this duty now become a reason for her to retroactively mail a check to the company, with this reason grounding an associated compensatory duty (60). Gardner calls this account of remedial duties the *continuity thesis*, and it seems to function as an argument for the compensation thesis: *A* acquires a duty to compensate *C* in virtue of her ϕ-ing because the ground of her duty not to ϕ becomes a duty-generating reason to compensate *C*.

While this is an elegant explanation of *A*'s remedial duty, it is vulnerable to two objections. First, the extent to which the continuity argument supports the compensation thesis depends on the specific reasons that ground all existing and future duties. In the bus case, it is apparent how post-infringement compensation advances the same end as discharging the initial duty; however, it is far from clear that the ground of every duty will also be a reason for the duty

51 Thanks to an anonymous reviewer for pressing me to be clearer about the role that empirical hypotheses play in these arguments.

52 Weinrib, "The Gains and Losses of Corrective Justice," 295; Ripstein, "As If It Had Never Happened," 1979; and Gardner, *Torts and Other Wrongs*.

53 All parenthetical citations are to page numbers in Gardner, *Torts and Other Wrongs*.

holder to provide compensation after she breaches that duty. If it turns out that there are even some duty-grounding reasons such that one cannot conform to those reasons by providing compensation, then the continuity argument will not support the compensation thesis in its very general form.

Second, even if one sets aside this first objection, the continuity argument does not support the compensation thesis's contention that *A uniquely* has a duty to compensate *C* in virtue of her ϕ-ing. To see this, consider an abstract version of Gardner's bus case wherein everyone has a duty not to ϕ grounded in some shared reason *R*.[54] Further, suppose that *A* breaches her duty/fails to conform to *R* by ϕ-ing, where the next best mode of conformance is compensating *C* by ψ-ing. Thus, according to the continuity thesis, *R* now grounds *A* having a compensatory duty to ψ. But why does third party *B* not also have a duty to ψ? After all, *B* had the same duty not to ϕ as *A*, where that duty was grounded in *R*. And now, just like *A*, she cannot conform to *R* by making it such that *C* is not subjected to ϕ-ing. Finally, given that *A*'s optimal conformance with *R* requires ψ-ing under such circumstances, it seems that conformance with *R* would equally require that *B* ψ under the circumstances—where this requirement gives rise to a remedial duty to ψ.

In short, *A* and *B* seem to be symmetrically situated vis-à-vis *R*, ϕ, and ψ. Thus, if the continuity thesis implies that *A* acquires a remedial duty to ψ, it seemingly also must imply that *B* acquires a duty to ψ. Or, put another way, Gardner seems committed to affirming that when it comes to conforming to *R*, the best thing to do is ensure that *C* does not suffer from ϕ-ing and, failing that, ensuring that *C* is the beneficiary of an act of ψ-ing. Given that by hypothesis, *R* applies equally to both *A* and *B*, *B* must ψ to conform with her reasons. Thus, the continuity thesis seemingly implies that *B* also has a remedial duty to ψ. Such a result fails to support the compensation thesis's contention that *A*—and *A* alone—has a duty to ψ given her ϕ-ing.

This point can be made less abstract by applying it to Gardner's bus case. There it was posited that each person has a duty to pay her fare because (a) she has reason to ensure that the bus company continues to operate, and (b) paying the fare conforms to that reason—presumably because, absent payment, the company will scale back its operations. When *A* breaches her duty by failing to pay her fare, she threatens the operations of the bus company; thus, Gardner

54 Some might wish to individuate actions in a more fine-grained way such that only a particular agent can ϕ because ϕ-ing incorporates her body and no one else's. For simplicity, it will be assumed that actions are individuated in a slightly more coarse-grained way such that different persons can ϕ; however, for those attracted to fine-grained action individuation, the case can be redescribed in terms of *R* being a reason for each person to undertake her own agent-relative action to realize a particular non-agent-relative state of affairs.

concludes that she must pay compensation to conform to her reason, as such compensation will have a similar effect with respect to sustaining the company's operations. However, note that some other bus rider *B* equally has the same reason to pay *her* fare—namely, she aims to preserve the company's operations. And now that *A* has threatened these operations via nonpayment, it seems that *B* too has reason to mail a compensatory check to the company to keep the buses running (with her payment being just as good as *A*'s payment vis-à-vis sustaining the company's operations). Granted, such action might not be needed if she knew that *A* were mailing a check instead. And she would be better off if *A* were the one to compensate the company. However, the same things can be said of *A* with respect to *B*. Thus, as far as having reason to compensate goes, the two seem to be symmetrically situated, with the continuity argument failing to demonstrate that *A*—rather than *B*—has a duty to provide compensation.

Against this objection, the proponent of the continuity argument might reply that *A* and *B* are not in fact symmetrically situated in either this case or its more generalized version above. This is because *A*, unlike *B*, has previously failed to conform to *R* by ϕ-ing, where any adequate interpretation of the continuity thesis holds that this gives *A* special reason to compensate *C* such that only *A* ends up with a compensatory duty. The problem with this reply is that it begs the question. Note that the proposition that *A* has a special duty-grounding reason to compensate *C* in virtue of breaching her duty not to ϕ is just a reassertion of the compensation thesis. Thus, if the continuity thesis is interpreted in these terms—or the continuity argument relies upon this proposition in any way to establish the asymmetry of *A* and *B* vis-à-vis *C*, *R*, and ψ-ing, then it is assuming what it aims to prove, namely, that prior duty infringement grounds a unique duty to compensate.

There may well be further arguments for the compensation thesis beyond the six listed above. Given this possibility, this section cannot be treated as a conclusive demonstration that no argument for the thesis succeeds. That said, there is, *prima facie*, limited reason for endorsing the compensation thesis. It should thus be rejected to both resolve the trilemma of section 1 and avoid any contradiction with rights-based luck egalitarianism.

9. CONCLUSION

This paper has presented two arguments against the compensation thesis. First, it has argued that the thesis is incompatible with (a plausible interpretation of) luck egalitarianism, as the latter posits that blameworthiness is a necessary condition of forfeiting rights to advantage while the former implies that rights to advantage are forfeited in cases where that condition is not met. Second,

the paper has presented a trilemma for rights theorists: in cases of blameless wronging, one cannot simultaneously hold that (1) someone is wronged, (2) the compensation thesis is true, and (3) it is unfair for a blameless party to have to fully internalize the costs of her action, where this unfairness renders a posited theory unacceptable. The paper has argued that one cannot resolve this trilemma by rejecting either proposition 1 or proposition 3. Thus, one must resolve the trilemma by rejecting the compensation thesis. Finally, the paper has objected to six potential arguments in support of the thesis. Given the absence of a successful supporting argument for the compensation thesis and the two arguments against it, the compensation thesis should be rejected.[55]

Te Herenga Waka—Victoria University of Wellington
jesse.e.spafford@gmail.com

REFERENCES

Alexander, Larry. "The Moral Magic of Consent (II)." *Legal Theory* 2, no. 2 (1996): 165–74.

———. "The Ontology of Consent." *Analytic Philosophy* 55, no. 1 (2014): 102–13.

Arneson, Richard. "Liberalism, Capitalism, and 'Socialist' Principles." *Social Philosophy and Policy* 28, no. 2 (2011): 232–61.

Cohen, G. A. *Rescuing Justice and Equality*. Harvard University Press, 2009.

———. *Why Not Socialism?* Princeton University Press, 2008.

Coleman, Jules. *The Practice of Principle: In Defence of a Pragmatist Approach to Legal Theory*. Oxford University Press, 2001.

———. *Risks and Wrongs*. Cambridge University Press, 1992.

Davis, Nancy. "Rights, Permission, and Compensation." *Philosophy and Public Affairs* 14, no. 4 (1985): 374–84.

Dougherty, Tom. *The Scope of Consent*. Oxford University Press, 2021.

———. "Yes Means Yes: Consent as Communication." *Philosophy and Public Affairs* 43, no. 3 (2015): 224–53.

Dworkin, Ronald. *Sovereign Virtue*. Harvard University Press, 2000.

Feinberg, Joel. "Voluntary Euthanasia and the Inalienable Right to Life." *Philosophy and Public Affairs* 7, no. 2 (1978): 93–128.

Frederick, Danny. "*Pro Tanto* Versus Absolute Rights." *Philosophical Forum* 45,

55 I am grateful to Adina Preda for her many helpful comments on an early version of this paper. The paper also benefitted from the comments of two anonymous reviewers, as well as from the questions and comments raised by audience members at the Te Herenga Waka—Victoria University of Wellington philosophy program colloquium.

no. 4 (2014): 375–94.
Gardner, John. *Torts and Other Wrongs*. Oxford University Press, 2019.
Goldberg, John C. P., and Benjamin C. Zipursky. *Recognizing Wrongs*. Harvard University Press, 2020.
Gordon-Solmon, Kerah. "What Makes a Person Liable to Defensive Harm?" *Philosophy and Phenomenological Research* 97, no. 3 (2018): 543–67.
Hohfeld, Wesley. "Some Fundamental Legal Conceptions as Applied in Judicial Reasoning." *Yale Law Journal* 23, no. 1 (1913): 16–59.
Hurd, Heidi M. "The Moral Magic of Consent." *Legal Theory* 2, no. 2 (1996): 121–46.
Hurley, Susan. *Justice, Luck, and Knowledge*. Harvard University Press, 2003.
Jorgensen Bolinger, Renée. "The Moral Grounds of Reasonably Mistaken Self-Defense." *Philosophy and Phenomenological Research* 103, no. 1 (2021): 140–56.
———. "Moral Risk and Communicating Consent." *Philosophy and Public Affairs* 47, no. 2 (2019): 179–207.
Knight, Carl. "An Argument for All-Luck Egalitarianism." *Philosophy and Public Affairs* 49, no. 4 (2021): 350–78.
———. "Egalitarian Justice and Expected Value." *Ethical Theory and Moral Practice* 16, no. 5 (2013): 1061–73.
Lang, Gerald. "How Interesting Is the 'Boring Problem' for Luck Egalitarianism?" *Philosophy and Phenomenological Research* 91, no. 3 (2015): 698–722.
———. *Strokes of Luck: A Study in Moral and Political Philosophy*. Oxford University Press, 2021.
Lippert-Rasmussen, Kasper. *Luck Egalitarianism*. Bloomsbury Publishing, 2015.
Lomasky, Loren. "Compensation and the Bounds of Rights." In *Compensatory Justice: Nomos XXXIII*, edited by John W. Chapman. New York University Press, 1991.
———. *Rights Angles*. Oxford University Press, 2016.
McGreer, Vitoria, and Philip Pettit. "The Hard Problem of Responsibility." In *Oxford Studies in Agency and Responsibility*, vol. 3, edited by David Shoemaker. Oxford University Press, 2015.
McGregor, Joan. *Is It Rape? On Acquaintance Rape and Taking Women's Consent Seriously*. Ashgate Publishing, 2005.
McKerlie, Dennis. "Equality and Time." *Ethics* 99, no. 3 (1989): 475–91.
———. "Justice Between the Old and the Young." *Philosophy and Public Affairs* 30, no. 2 (2001): 152–77.
McMahan, Jeff. "The Basis of Moral Liability to Defensive Killing." *Philosophical Issues* 15 (2005): 386–405.
———. *Killing in War*. Oxford University Press, 2009.

Miller, David. "Justice." In *Stanford Encyclopedia of Philosophy* (Fall 2023). https://plato.stanford.edu/archives/fall2023/entries/justice/.

Montague, Phillip. "Davis and Westen on Rights and Compensation." *Philosophy and Public Affairs* 14, no. 4 (1985): 390–96.

Olsaretti, Serena. "Responsibility and the Consequences of Choice." *Proceedings of the Aristotelian Society* 109, no. 2 (2009): 165–88.

Otsuka, Michael. "Killing the Innocent in Self-Defense." *Philosophy and Public Affairs* 23, no. 1 (1994): 74–94.

Parfit, Derek. "Equality or Priority?" In *The Ideal of Equality*, edited by Mathew Clayton and Andrew Williams. Palgrave Macmillan, 2000.

Preda, Adina. "Are There Any Conflicts of Rights?" *Ethical Theory and Moral Practice* 18, no. 4 (2015): 677–90.

Quong, Jonathan. "Rights Against Harm." *Proceedings of the Aristotelian Society, Supplementary Volumes* 89 (2015): 249–66.

Ripstein, Arthur. "As if It Had Never Happened." *William and Mary Law Review* 48, no. 5 (2007): 1957–97.

———. *Private Wrongs*. Harvard University Press, 2016.

Segall, Shlomi. *Why Inequality Matters: Luck Egalitarianism, Its Meaning and Value*. Cambridge University Press, 2016.

Shafer-Landau, Russ. "Specifying Absolute Rights." *Arizona Law Review* 37, no. 1 (1995): 209–24.

Skorupski, John. *The Domain of Reasons*. Oxford University Press, 2010.

Spafford, Jesse. "Luck Egalitarianism Without Moral Tyranny." *Philosophical Studies* 179, no. 2 (2022): 469–93.

———. Review of *Strokes of Luck: A Study in Moral and Political Philosophy* by Gerald Lang. *Ethics* 133, no. 3 (2023): 429–34.

———. *Social Anarchism and the Rejection of Moral Tyranny*. Cambridge University Press, 2023.

Tadros, Victor. "Causation, Culpability, and Liability." In *The Ethics of Self-Defense*, edited by Christian Coons and Michael Weber. Oxford University Press, 2016.

Temkin, Larry. *Inequality*. Oxford University Press, 1993.

Thaysen, Jens Damgaard, and Andreas Albertsen. "When Bad Things Happen to Good People: Luck Egalitarianism and Costly Rescues." *Politics, Philosophy and Economics* 16, no. 1 (2017): 93–112.

Thomson, Judith Jarvis. *The Realm of Rights*. Harvard University Press, 1990.

Vallentyne, Peter. "Brute Luck and Responsibility." *Politics, Philosophy and Economics* 7, no. 1 (2008): 57–80.

Weinrib, Ernest J. "The Gains and Losses of Corrective Justice." *Duke Law Journal* 44, no. 2 (1994): 277–97.

———. *The Idea of Private Law*. Oxford University Press, 2012.
Wellman, Christopher Heath. "On Conflicts Between Rights." *Law and Philosophy* 14, nos. 3–4 (1995): 271–95.
Wenar, Leif. "Rights." In *Stanford Encyclopedia of Philosophy* (Spring 2023). https://plato.stanford.edu/archives/spr2023/entries/rights/.

Journal of Ethics and Social Philosophy
VOL. 30, NO. 3 · MAY 2025
https://doi.org/10.26556/jesp.v30i3.3505

CHALLENGES FOR THE INABILITY THEORY OF DISABILITY

Stephanie Elsen

WHY IS IT IMPORTANT to think about what makes a condition a disability? Many of us aim to shape social practices in a way that pays more attention to the situation of people with disabilities in society, and we expect the same from the state's social systems. Doing so assumes that there is a basis on which we can justifiably classify certain conditions as disability. But at the same time, we do not seem to have a good grasp of what it means to have a disability. As a result, we are unsure whose situation needs to be considered when it comes to the topic of disability. There is also the question of whether it is at all meaningful to think about moral issues specifically in relation to disability. For example, instead of thinking more specifically about the situation of people with disabilities who are affected by loneliness, why not think exclusively about what obligations the situation of *all* those affected by loneliness imposes on us? Justifying the need for such a special focus in moral reasoning requires an account of what makes something a disability.

In his 2020 paper "Disability as Inability," Alex Gregory offers a disability account that promises to fill the explanatory gap in our classificatory, social, and moral practices related to disability.[1] On his view, what distinguishes cases of disability from cases of nondisability is a particular kind of *inability* that is determined relative to the statistically most common body features and abilities in a reference class. It is this particular kind of inability that is shared by individual disabilities such as "motor neurone disease, diabetes, achondroplasia, deafness, a missing lower arm, and cerebral palsy" (23).

The inability theory ties in with two aspects of how we commonly use the term 'disability' in everyday life and therefore appears attractively simple and intuitive. These are the ideas that disability is about a limitation of abilities and a deviation from the typical or normal. There are already a number of proposals that take up these two features of our language use, but these proposals face

1 Gregory, "Disability as Inability" (hereafter cited parenthetically). Gregory has these aims in mind when he motivates and defends the inability theory (23, 24). In addition, there are reasons to believe that the inability theory is a promising starting point for addressing these challenges. I come back to this point below.

intricate difficulties.[2] The inability theory is an attempt to follow this natural line of thought in a way that avoids the problems of earlier versions.[3]

This paper starts from the assumptions that we need a disability account and that the inability theory is in many respects a convincing proposal for such an account. The main focus, however, will be the features of the inability theory that have costly disadvantages and therefore motivate the search for a disability account that can do without these features.

I proceed as follows. In section 1, I introduce Gregory's adequacy criteria for a disability account and his proposal for meeting these criteria—*the inability theory of disability*. Further, I present how he rejects common objections to the classificatory adequacy of the theory and what he sees as the main merits of his proposal. In section 2, I elaborate on some of the objections that have already been raised against Gregory's proposal or, more generally, against disability accounts that focus on ability limitations. I use four types of examples to show that Gregory's theory continues to struggle with problems of over- and under-inclusiveness, or fails to provide guidance for classifying disability. In section 3, I raise a more fundamental problem for the inability theory. I argue that it meets one of Gregory's central adequacy criteria for a disability account at the cost of not meeting the other. I conclude in section 4.

1. THE INABILITY THEORY

With the inability theory, Gregory pursues two goals:

> I aim to provide an account of disability that picks out something real in the world—to offer a theory that carves nature at the joints. Such a descriptive theory would show what various disabilities have in common and explain what distinguishes them from other superficially similar phenomena.... In fact, we need a theory that carves nature at the joints *and* whose content is tolerably close to our everyday concept of disability so that we can recognize the theory as a more careful and complete version of that everyday concept. (24, 25)

2 Other disability accounts that assign a crucial role to inabilities include Amundson, "Disability, Handicap, and the Environment"; Hull, "Defining Disability"; World Health Organization, "International Classification of Impairments, Disabilities, and Handicaps," 143, and "International Classification of Functioning, Disability and Health," 213; Buchanan et al., *From Chance to Choice*, 286; and Begon, "Disability."

3 The way David Wasserman and Sean Aas include Gregory's proposal in their contribution "Disability" in the *Stanford Encyclopedia of Philosophy* suggests that Gregory's inability theory is being received along these lines.

His objective of developing an account that picks out something real in the world we might call the *realist criterion*. This criterion ensures that a disability account provides a standard of correctness that does not merely derive from our intuitive everyday classifications of disability. But by claiming that the theory's content should be "tolerably close to our everyday concept of disability," Gregory accepts a further adequacy criterion. We might call this one the *conservative criterion*. It says that cases of disability and nondisability considered paradigmatic in everyday judgments about disability are to be classified accordingly by the theory.

According to Gregory's final version, the inability theory says:

> To be disabled is to be less able to do something than is typical, where this degree of inability (1) is partly explained by features of your body that are atypical and (2) is not explained by anyone's attitudes toward those bodily features. (33)

On Gregory's view, whether body features are atypical depends on which body features are statistically most common among people of the same sex and stage of development (for example, among female adults). Atypical inabilities are defined analogously. Whether the relevant kind of inability is present depends on which abilities and ability levels are statistically most common among people of the same sex and stage of development.

The inability theory allows that abilities are individuated in such a way that they can also be capacities to carry out relatively specific actions or activities, such as "playing a piano *with your fingers*" (27). Furthermore, according to Gregory, an inability can be present both when someone is completely lacking an ability or when someone has the ability but to a lesser degree. By means of this latter feature, the inability theory avoids one familiar objection, which says that disabilities—such as conditions that affect energy levels or achondroplasia—often do not prevent one from doing anything specific and are very variable in their effects.[4] On the inability theory, even if such conditions do not generally prevent one from carrying out a particular activity, they may have the effect that one can no longer carry out a particular activity for a certain period of time, such as working a forty-two-hour week (28).

Gregory also assumes that there is no sharp line between statistically typical and statistically nontypical bodily features or ability levels (29). In this way, he also addresses another objection, which claims that the inability theory is overinclusive. Someone can have slightly less typical body characteristics—"as

4 Barnes, *The Minority Body*, 17.

being a petite woman"—and still be within the range of the typical and therefore not count as having a disability, despite being less able in some dimensions.[5]

A similar objection aimed at overinclusiveness states that on the inability theory, the lack of statistically typical but intuitively completely irrelevant abilities such as tongue rolling and ear wiggling also counts as disability.[6] Gregory's response is to accept this result and instead explain away our intuitions that the lack of these abilities is not a disability. His explanation is that, strictly speaking, these missing abilities are disabilities, but to actually *describe* them as disabilities would be highly misleading because we usually describe something as a disability only when the condition is significant and practically relevant. In the context of such a speech act, one would say nothing false but something nevertheless inappropriate because it violates the generally applicable conversation maxim of relation, as specified by Grice.[7]

Another familiar objection to (in)ability theories of disability in general refers to persons with bodily characteristics that are statistically atypical (such as a certain skin or hair color) who are directly discriminated against because of these characteristics and who are restricted in their abilities as a result.[8] It is because of examples of this type that Gregory introduces condition 2, according to which only those inabilities that are not explained by the attitudes of others towards these bodily features are disabilities (33). Without this condition, someone who is not able to leave the house without restrictions due only to a certain skin color and attendant social prejudices against people with this skin color would count as disabled.

Gregory points to several important merits of his view. I present three of them here. First, he argues that the inability theory makes clear what all those apparently different conditions that we typically classify as disability have in common and thereby also provides "the most natural explanation of why we categorize these things together" (26). Second, he argues that the inability theory identifies a feature that also explains why disabilities are considered politically relevant. The inability theory makes it intelligible why we specifically think about requirements of justice with respect to people with disabilities and have social practices organized around the category of disability (26, 27,

5 The objection Gregory responds to in this way is raised by Elizabeth Barnes using the example of the petite woman against disability accounts that assign a central role to the lack of abilities (Barnes, *The Minority Body*, 17, 20).

6 Barnes, *The Minority Body*, 16; and Kahane and Savulescu, "The Welfarist Account of Disability," 45.

7 Grice, "Logic and Conversation," 46.

8 Barnes, *The Minority Body*, 19.

36). Third, he stresses that the inability theory provides us with a criterion that allows us to categorize unclear cases such as nut allergies (27).

There is another feature of the inability theory that is worth highlighting. The view does justice to the widespread belief that disability is normatively relevant without having to define disability as something bad or suboptimal.[9] As Gregory points out, the extent of our abilities seems to influence our amount of freedom and thereby closely linked to something to which we often ascribe a particular normative weight (35, 36). So it is certainly not far-fetched to discuss disability, understood as a particular form of inability, as something normatively relevant. At the same time, Gregory's definition does not include normative vocabulary.[10] Rather, it remains a separate question how different ways of "being less able to do something than is typical" are related to individual well-being levels.

If Gregory's inability theory really has all the advantages mentioned above, then it could be successfully used as a basis for discussing the adequacy of our classificatory and social practices, and for raising moral questions that particularly relate to people with disabilities. Unfortunately, there are reasons to doubt that the inability theory, in its present form, actually has all the appealing features mentioned above. In the next two sections, the focus is on classificatory problems for Gregory's account that go beyond those that have already been raised by others, as well as on a more fundamental problem arising from the fact that Gregory accepts both the conservative criterion and the realist criterion.

2. The Inability Theory Captures Too Much and Too Little

The inability theory has difficulties in defining a criterion for the relevant inabilities and bodily characteristics that (a) neither excludes paradigmatic cases of disability nor includes paradigmatic cases of nondisability and (b) gives a clear answer as to how specific cases are to be classified. I use certain cases of visual limitations that occur in a particular context as examples of paradigmatic cases of disability that the inability theory cannot classify as cases of disability, and these cases thus serve as counterexamples to the theory. As paradigmatic

9 Elizabeth Barnes presents the stronger claim that not defining disability as "something that's bad or suboptimal" is a criterion of success for a disability theory rather than merely an attractive feature (*The Minority Body*, 11).

10 In this respect, Gregory's inability theory differs from another recent proposal to understand disability in terms of a limitation of abilities. Jessica Begon narrows down the ability limitations relevant to disability by saying that they involve "restriction in the ability to perform those tasks human beings are entitled to be able to perform as a matter of justice" ("Disability," 936, 937).

examples of nondisability, I present left-handedness, menopause, and pregnancy as cases that the inability theory cannot clearly classify correctly. As we will see, these counterexamples arise because Gregory's account is so far elaborated only in broad outlines and because statistical standards play a central role in it.

Importantly, the aim of this section is not to point to uncontroversial and insurmountable tensions between classificatory judgements based on the inability theory and our actual classification practice. The aim in outlining the classificatory problems for Gregory's account is to emphasize or draw attention to particular features of the inability theory that make it significantly more difficult to arrive at clear, justified, and—in terms of our classification practice—convincing judgements about specific cases. Thus, even if readers are not convinced that all the examples discussed are paradigmatic cases of disability/nondisability and provide clear counterexamples to the inability theory, they can take the following discussion as a way of demonstrating the extent to which certain features of the inability theory pose challenges to this proposal that are not easily met.

From the discussion of certain visual impairments in a specific context and left-handedness, we learn that the inability theory provides us with classification criteria that do not allow clear classifications, at least in certain cases, because the theory does not include a definition of the range of the statistically atypical bodily characteristics relevant to disability.[11] An interesting aspect of these examples is that the classification of the cases based on the inability theory remains indeterminate even when we modify the details of the cases in ways that tend to influence our pretheoretical judgements about these cases. The discussion of left-handedness further illustrates that we can easily come up with statistically atypical ability limitations for statistically atypical bodily characteristics. This type of example suggests that the inability theory may classify many more things as disabilities than we usually assume, and Gregory's Gricean strategy therefore needs to be applied much more often than it might seem in Gregory's discussion of a similar overinclusiveness objection from the literature.[12] The examples of menopause and pregnancy demonstrate the extent to which disability classifications based on the inability theory depend on the definitions of the relevant reference classes and thus on convincing explanations of why certain ways to define those reference classes are more adequate than others.

11 Gregory addresses this concern, but not to the extent necessary to refute it ("Disability as Inability," 29). See note 16 below for further discussion.

12 Kahane and Savulescu, "The Welfarist Account of Disability," 45.

I start with an objection of underinclusiveness. There are conditions that intuitively constitute disabilities, at least in certain contexts, but that the inability theory cannot classify as such because the ability limitation is caused by physical characteristics that are *comparatively typical*. Examples are ametropias such as short- or long-sightedness, astigmatism, and presbyopia. These are widespread but nevertheless seem to constitute disabilities in certain contexts. Presbyopia, for example, is defined "as the gradual and progressive age-related loss of accommodative amplitude and is ultimately due to an age-related loss in the ability of the lens to undergo accommodative optical changes."[13] For typically sighted individuals, presbyopia is associated with "blurred vision at near, visual fatigue or headache after attempting to read at near for prolonged periods, or an inability to sustain clear vision at a normal reading distance" from an age of forty to forty-five years. Since presbyopia is age related and because of our long lifespan, its symptoms are very common.

At first glance, the inability theory appears to have no difficulty in attributing disability in presbyopia cases. Gregory could say that presbyopia consists of atypical bodily features and is associated with atypical levels of ability (partly because of these bodily features) and that the conditions for the presence of a disability are therefore met. On the inability theory, typicality depends on reference groups that are also determined by one's stage of development. And a stage of development such as being biologically adult is a very broad category, as Gregory understands it. It includes a thirty-year-old as well as an eighty-year-old (28). Relative to such a large reference group, the bodily condition constituting presbyopia and its later symptoms most likely do not qualify as statistically most frequent and thus as typical in this strong sense.

However, Gregory acknowledges a range of typical bodily characteristics and ability levels. This means that he allows for less typical bodily features and ability levels that are still not atypical in the way constitutive of disability. Recall the example of "a petite woman," who has slightly less typical bodily characteristics that also diminish her ability levels in some dimensions. To reject this case as a counterexample to the inability theory, Gregory argues, "But whilst petite women might have somewhat atypical bodies and thereby lack some relatively typical abilities, their bodies and ranges of ability are not *that* atypical" (28). The first concern about the inability theory in its current state is now that as long as Gregory does not more clearly outline the range of the typical, it remains unclear why the same that holds for the example of "petite women" should not also apply to presbyopia. Presbyopia and its symptoms are *somewhat* atypical, not *very* atypical, and therefore not atypical to the extent

13 Glasser, "Presbyopia," 489.

constitutive of disability. Thus, since Gregory allows for a range of the typical, it does not seem controversial to say that his account places presbyopia and its symptoms as within the range of typical bodily characteristics and thus does not classify presbyopia and its symptoms as a disability.

But is this really a problem for the inability theory? Are we indeed facing a paradigmatic case of disability that cannot be adequately classified by the inability theory? What seems to indicate that we are not dealing with a paradigmatic case of disability is that presbyopia appears to be associated with only a slight reduction in typical vision, which can be easily corrected with a visual aid. My concern with the inability theory becomes obvious when we consider that we want to come up with a disability account that can be applied globally and that also considers life circumstances that are very different from those of the typical authors of academic papers. Just think of a person with presbyopia who does not have access to visual aids and earns her living by sewing.[14] According to our everyday judgments about disability, people with presbyopia clearly have a disability, at least under these circumstances, and Gregory cares about our everyday judgments. This case highlights that whether we consider something to be a paradigmatic case of disability seems to depend also on the environment. When we consider an environment where a person has no access to visual aids but relies on very good vision, we are probably more inclined to judge that the person has a disability.

On a general level and in relation to other cases, the inability theory can also address the importance of a particular environment for our classification practice. As Gregory notes, a change in the environment can make it the case that an atypical bodily characteristic no longer results in an atypical level of ability that constitutes a disability (30). However, addressing the importance of a particular environment in this way does not alter the fact that the inability theory cannot classify disability in accordance with our classification practice or provide guidance for our classification practice in cases of presbyopia. Also in the case of someone with presbyopia who has no access to visual aids yet relies on good vision, it is not clear whether the bodily characteristics that constitute presbyopia are within or outside the range of statistically typical bodily characteristics. As I wrote before, as long as Gregory does not more clearly outline the range of the typical, it remains unclear why the same that holds for the example of "petite women" should not also apply to presbyopia: presbyopia and its symptoms are somewhat atypical, not very atypical, and therefore not atypical to the extent constitutive of disability. And thus, even in the specific setting in which we are inclined to classify presbyopia as a paradigmatic case of

14 See Holden et al., "Towards Better Estimates of Uncorrected Presbyopia."

disability, it is unclear whether the necessary conditions for disability proposed by the inability theory are met. To sum up: since Gregory refers to statistical standards and does not further specify the range of what is statistically typical that is relevant to disability, the inability theory faces difficulties in classifying presbyopia as a disability, even in specific cases in which presbyopia seems to be a paradigmatic case of disability.

Another example indicating the underinclusiveness of the inability theory is a significant reduction of the ability to concentrate due to nutrient deficiencies or thyroid problems. In this type of case too, we are dealing with an ability limitation that is caused by comparatively typical bodily characteristics, and thus one of Gregory's necessary conditions for disability is not clearly met. As with the example of presbyopia, the extent to which we judge such cases to be counterexamples to the inability theory, thus demonstrating its underinclusiveness, depends on the context specific to individual cases. (For example, is it easy to get thyroid medication, and is it well tolerated?) But even if we assume that conditions such as nutritional deficiencies and thyroid problems are not paradigmatic cases of disability, not even in the scenarios where they are not compensated for with medication and supplements, these cases still show us that the inability theory provides us with classification criteria that often do not provide clear answers.

Let us now turn to concerns about overinclusiveness. We can start with generalizing the objection already mentioned above that on the inability theory, the lack of statistically typical but intuitively completely irrelevant abilities such as tongue rolling and ear wiggling also counts as disability. Since Gregory does not specify the inability theory in a way that excludes these cases, we can identify a statistically atypical ability level for virtually any statistically atypical physical characteristic.[15] As a result, when we deal with statistically atypical physical characteristics, we would also be dealing with disabilities. Since Gregory allows for very fine-grained descriptions of abilities and does not presuppose the complete absence of an ability for a disability to be present, there are no limits to our creativity in coming up with counterexamples. Gregory's proposal to explain away our divergent intuitions for such cases with reference to Grice's communication maxim of relevance would therefore have to be applied much more often. In the context of the characteristic of left-handedness, for example, we find many examples of reduced ability levels that might simply be related to the fact that most people are right-handed rather than to the fact that left-handed people are socially neglected or subject to discriminatory attitudes (yet such unjust reductions in ability levels are certainly part of our social reality too).

15 I am very grateful to Andreas Cassee for helping me to see this point.

And some of the reduced ability levels that affect left-handed people because of their handedness may have no relevance to them at all. In these cases, we would probably not assume any paradigmatic disabilities. However, based on the inability theory, the necessary and sufficient conditions for a disability would be met in these cases, which include irrelevant atypical ability levels partly due to left-handedness.

Besides the reference to Grice's communication maxims, Gregory could of course argue with regard to this case that left-handedness and the ability limitations associated with it are not *so* atypical. But as we have seen in the discussion of presbyopia, this strategy has the effect that other intuitive cases of disability and nondisability can no longer be captured by the theory of inability because they are based on physical characteristics that are not *so* atypical either.[16]

A second source for overinclusiveness problems is Gregory's characterization of relevant reference classes, in particular that he distinguishes them according to biological stages of development. As Gregory understands the developmental stage of adulthood, it includes people of very different ages, which are thus classified by the same typicality standards. It follows from this that "an eighty-year-old with inabilities that are typical for someone their age may nonetheless be disabled if those inabilities are not typical for human adults in general. In turn, the theory rightly entails that many elderly people are disabled" (29).

Gregory's theory captures something plausible but overshoots the mark. Mobility restrictions such as an insecure gait without aids due to very low muscle tone seem plausible examples of age-related disabilities. However, the situation is different with other physical characteristics such as menopause. Given the wide age range within the group of biological adults, menopause, which is associated with ability limitations, might likely be classified as a disability by the inability theory. But this clearly contradicts our everyday classifications, at least when we think of menopause as occurring after the age of forty.

Finally, there are examples of overinclusiveness that suggest that Gregory's definition of reference classes needs further types of restrictions, in addition to

16 I take it that an ability limitation due to being left-handed is a paradigmatic example of *not* having a disability and that, e.g., difficulties with sewing for a living due to presbyopia is a paradigmatic example of *having* a disability. As long as this is accepted, it is also clear that Gregory cannot in all cases take his theory's indeterminacy as to when something is atypical in the relevant sense as an advantage, as something that adequately captures the phenomenon of disability. See Gregory, "Disability as Inability," 29. And even if one does not accept that left-handedness is a paradigmatic case of nondisability, and presbyopia in a particular setting is a paradigmatic case of disability, these and analogous cases suggest that the inability theory is unable to provide a clear answer in relation to a large set of cases. This result does not fit well with the desideratum to develop a theory of disability that provides us with criteria for categorizing unclear cases. See Gregory, "Disability as Inability," 24, 27.

species, sex, and stage of development. Take the example of pregnancy. Pregnant women are confronted with a number of atypical bodily characteristics that lead to limitations in their ability levels, and yet we do not usually classify their condition as a disability. They gain a lot of weight in a short time, which leads to mobility restrictions. They are quicker to be out of breath, which influences their ability to pursue sports activities in the usual way. They have greater need for certain nutrients and face hormonal changes that can affect their energy balance. But if they are subject to the same standards as nonpregnant women, the inability theory may classify them as having a disability.[17] Another candidate for further restriction of reference classes is skin color. There is evidence that people with darker skin color who live in Europe have problems with vitamin D balance, while some people with lighter skin color living near the equator have problems with folic acid balance.[18] However, attributing disability to individuals who have nutritional problems due to a combination of their skin color and geographical location seems to run counter to our practice of classifying disability.

A natural way to address these problems of overinclusiveness is to further restrict the relevant reference classes. However, there are in principle many different ways to determine the relevant reference class, and the question arises why certain characterizations are more adequate than others.[19] This points us to a more fundamental challenge for the inability theory, stemming from the fact that Gregory accepts both the conservative criterion and the realist criterion, which is the topic of the next section.

3. Meeting Both the Conservative and the Realist Criteria

The inability theory requires reference to specific reference classes. Distinguishing between reference classes captures that what is typical for one group is not typical for another group. If the relevant inabilities were not determined relative to what Gregory terms "typical for a human being of your sex at your

17 I make the weaker claim that the inability theory *may* imply that pregnancy is a disability because—as I have discussed before—the inability theory is underdetermined as to when something is atypical in a disability relevant way. Because of this ambiguity, only limited claims can be made about what the inability theory implies for particular cases.

18 Harris, "Vitamin D and African Americans"; and Jones et al., "The Vitamin D-Folate Hypothesis as an Evolutionary Model for Skin Pigmentation."

19 For a presentation of this line of criticism against Boorse's biostatistical theory of health, see Kingma, "What Is It to Be Healthy?" and "Naturalist Accounts of Mental Disorder." Meanwhile, Wasserman and Aas have also drawn on Kingma's objection to Boorse's theory of health when discussing ability theories of disability, particularly Jessica Begon's proposal ("Disability").

stage of development," the inability theory would go against our intuitions that "humans are not disabled in virtue of lacking the ability to fly, men are not disabled in virtue of lacking the ability to breastfeed, and infants are not disabled in virtue of lacking the ability to talk" (28).

Further, to be tolerably close to our everyday concept of disability, the inability theory needs to include certain reference classes instead of others.[20] As we have seen in the previous section, it should probably distinguish between, for example, not only sexes but also pregnant and nonpregnant women. At the same time, the account should not allow different reference classes for wheelchair users and people who do not need a wheelchair. In this case, the condition of a person who needs a wheelchair would not be classified as a disability, because needing a wheelchair is not atypical compared to a reference class in which everyone needs a wheelchair. Hence it is not difficult to observe that regarding the objective to give an account that accommodates our everyday disability classifications, there are adequate and inadequate reference classes. The worry now is that the inability theory does not provide an account of what makes a reference class adequate and therefore cannot give us a satisfactory answer to what disability is.

Gregory does not explicitly justify his selection of reference classes. There are scant references to the underlying motivations for such choices. With reference to the key term 'typical', he writes, "'Typical' here means 'typical for a human being of your sex at your stage of development.'... (In principle, we might relativize further, such as to race. But it is hard to find *intuitive cases* that support further restrictions like this)" (28, emphasis added). The additional comment in parentheses suggests that Gregory takes everyday judgments about paradigmatic cases of disability or nondisability as a guideline for his choice of adequate reference classes. The reasoning seems to be that since we commonly would not judge that men are disabled in virtue of lacking the ability to breastfeed, distinguishing between reference classes according to sex is adequate. Evidently, this is the most obvious way to satisfy the conservative criterion. And to rely on everyday judgments about cases of disability does not seem to be, in principle, inconsistent with the second adequacy criterion to provide an account that picks out something real. Philosophers often look at everyday judgments about instances of *X* precisely because they hope to arrive at new ideas about what *X* actually is. The idea is that we may already be on the right track with our everyday judgments, that they capture something real, and that we can thus learn something from studying them. Moreover, Gregory refers only to everyday classifications

20 Here I am mirroring Kingma's line of criticism against Boorse's theory of health (Kingma, "What Is It to Be Healthy?" 128, 129).

of paradigmatic cases of disability and nondisability, and it is a widely accepted strategy to consider everyday classifications as part of a reflective equilibrium approach in the justification of an account. And yet the sole reference to our everyday classifications in the justification of the relevant reference classes raises questions when it comes to an account of disability.

The conservative and the realist criteria are in tension with each other because the latter, as a standard of correctness, provides a corrective to our everyday classifications when these are flawed. I argue that in the case of disability, an account's potential to provide a standard of correctness for our existing classifications is significantly weakened if key components of this account are designed only to map our existing classification practices.

In the case of disability, there seems to be a particular danger that our existing classifications do not carve nature at the joints, which makes the classifications inadequate to solely guide our theorizing about disability and makes a standard of correctness for our disability classifications all the more important. This danger relates to facts about the history and practice of classifying disability. First, it has been suggested that our existing classifications of disability are the result of complex interactive processes between many interest parties against the background of major historical events and changing political, legal, economic, and social circumstances.[21] The term 'disability' entered official, technical, and everyday language, and its meaning changed because it served the interests of varying groups.[22] Our everyday classifications thus seem not only variable but also opportunistically shaped. Second, our classifications have enormous social consequences. There is a lot at stake if disability is incorrectly classified. Such misclassification can be decisive for whether a person has access to important resources, what standing they enjoy in their social environment, whether their needs are adequately considered, and what others owe them. This social and normative dimension of disability classifications can also invite misclassification under certain circumstances. This possibility must be considered when a disability account is based on existing classifications. Third, disability classifications often affect individuals who do not themselves shape these classification conventions. These people do not have the capacities or the necessary external resources to actively participate in the practice of classification in the light of their experiences. Finally, there is much to suggest that our existing classifications of disability are also shaped by stereotypes. These include the idea that an individual with a disability can only achieve something

21 Linton, *Claiming Disability*; Silvers, "On the Possibility and Desirability of Constructing a Neutral Conception of Disability"; and Burch and Sutherland, "Who's Not Yet Here?"

22 Francis and Silvers, "Perspectives on the Meaning of 'Disability.'"

if she "overcomes" her disability or that a disability typically manifests itself in someone needing a wheelchair to get around.[23]

The more fundamental problem with the inability theory, then, is that a central component—the definition of the relevant reference classes—is based exclusively on conservative considerations. This means that a particularly error-prone classification practice is built into a theory that is supposed to capture what disability really is and could thus serve as a standard of correctness.

The inability theory has a second central component that is motivated by considerations of compliance with our everyday classifications. Again, the question arises as to whether the inability theory can fulfill the realist criterion in view of this type of motivation, considering how error prone our classification practice seems to be in the case of disability.

This second central component is condition 2 of Gregory's definition of disability. Let us recall the final version of the inability theory:

> To be disabled is to be less able to do something than is typical, where this degree of inability (1) is partly explained by features of your body that are atypical and (2) is not explained by anyone's attitudes toward those bodily features. (33)

According to Gregory, the second condition excludes only cases in which an individual's reduced level of ability results from the problematic attitudes of others towards that individual's atypical bodily characteristics. Not excluded are cases of reduced ability levels that result from living with atypical bodily characteristics in a social environment shaped by an unjust lack of attention to the situation of people with these bodily characteristics (33). Gregory presents the following scenario as an example of the first type of case, which makes condition 2 necessary in order to avoid counterexamples to the inability theory:

> Imagine that you are a member of a small minority race and are a victim of direct discrimination on the basis of your race. This racism might reduce the options you have. To that extent, you might be unable to do certain things, where this inability is partly explained by the atypical features of your body. So it seems as though our theory classifies you as disabled. But plausibly, under these circumstances, you are not necessarily disabled. (32)

23 For example, we often seem to take it for granted that a disability is visible and manifests itself in the use of mobility aids. Evidence of this is provided by the reports of people with so-called invisible disabilities, who often have to make special efforts to convince others of having a disability. See Stone, "Reactions to Invisible Disability."

Gregory wants the inability theory to conclude that people who face ability limitations due to direct discrimination based on, for example, their atypical skin color do not have a disability. The aim of avoiding such counterexamples, and thus following our classification practice, motivates condition 2.

My concern with condition 2 arises from the fact that for including 2, unlike 1, Gregory seems to rely only on reasons that are based on conservative considerations. Furthermore, unlike in the case of condition 1, it is unclear what a more comprehensive motivation for condition 2 might look like. What reasons other than conservative ones are there for accepting condition 2? Rather, there seem to be reasons against adopting condition 2 as part of a disability account.

A number of considerations speak in favor of making individual physical and mental characteristics, as referred to in condition 1, a central element of a disability account. What is crucial for the argument here is that these considerations do not exclusively concern our everyday classifications of disability. If one is interested in the question of what disability *is*, not just in finding a pragmatic answer that serves, for example, certain political, social, or administrative purposes, then it seems plausible to advocate an account according to which disability is not exclusively related to how one is treated or how one self-identifies but also has something to do with one's physical and mental characteristics.[24] Here are some reasons, which are not exclusively based on conservative considerations, for the inclusion of this individual component. First, various disability accounts already explain disability by reference to individual physical and mental characteristics.[25] Second, when we look at paradigmatic cases of disability, the specification of individual physical or mental characteristics typically plays an important role in the description of the respective situation. Third, many people with disabilities also emphasize the individual, physical, or mental side of disability, which does not mean that they see this side as something inevitably bad.[26] Fourth, there may be pragmatic reasons for explaining disability only in terms of, for example, certain attitudes in society, but from a scientific point of view, there are a whole range of different factors, including individual physical and mental characteristics, that interact with each other and affect an individual's ability level. All of these need to be taken into account in a disability account that is not subjective or value laden.[27]

24 Barnes, *The Minority Body*, 36–38.

25 Radical versions of the social model are an exception (e.g., Oliver, *Understanding Disability*). However, it is unclear whether these versions would even subscribe to the realist criterion and not just see themselves as tools to implement certain social and political objectives.

26 For an exemplary statement of this kind, see Clare, "Stolen Bodies, Reclaimed Bodies," 359.

27 Wasserman, "Philosophical Issues in the Definition and Social Response to Disability," 225–29.

To see the challenge for the inability theory, it is now crucial that we do not have such a variety of reasons in favor of condition 2. Furthermore, we have reasons that are not based on our everyday disability classifications against adopting condition 2. Statements by people with disabilities about their situations with disability often refer to the interaction of individual factors with environmental factors and, in particular, with the attitudes of other people.[28] There is evidence that the attitudes of fellow human beings to certain individual physical and mental characteristics have a significant influence on the situation of people with these characteristics, their opportunities, and perceived well-being:

> People with disabilities are believed to be incapable, useless, pitiful or even laughable. Stigma shapes the affect, attitudes and behaviour of others that mar the daily lives of people with disabilities: the profound condescension implied in being labelled an inspiration for performing ordinary tasks; being robbed of decision-making authority over matters of intimate personal concern; being mocked and ridiculed by colleagues, neighbours and strangers.[29]

Since the degrading attitudes of others are so central to the experience of disability, the question becomes even more pressing as to how condition 2 is justified, which explicitly excludes inabilities due to attitudes in response to atypical bodily characteristics from the account. Gregory does not have to deny that these attitudes have a significant influence on the situation of people with disabilities; he must deny only that the attitudes are relevant to the question of whether or not a persons has a disability. However, on the basis of the above considerations, it seems justified to expect Gregory to make an argument for condition 2 that is not exclusively based on facts about our classificatory practice: first, because of the role that attitudinal barriers play in the lives of many people with disabilities; second, because individual factors (as specified in condition 1) typically affect ability levels not in isolation but rather in complex interaction with a wide range of environmental and social factors.[30] Focusing on only certain types of factors in this complex interactive relationship requires a justification that also considers the dangers of our existing disability classification practices. Thirdly, a more comprehensive justification for condition 2 is required because we have reason to believe that our disability classifications

28 See, e.g., Young, "I'm Not Your Inspiration"; and Stock-Landis, "The Toxic Myths I Internalized as a Person with Facial Differences."

29 Barclay, *Disability with Dignity*, 129.

30 Wasserman, "Philosophical Issues in the Definition and Social Response to Disability," 225–29.

are particularly prone to error, and the stakes for those who might be affected by error are high.

The strong focus on conservative considerations is also evident in the way Gregory responds to a possible objection to condition 2. The objection is that the inability theory is too individualistic an approach to disability because it excludes a central group of attitudes towards individual bodily characteristics as disability-constituting features. To counter this objection, Gregory introduces the aforementioned distinction between inabilities resulting from certain problematic attitudes in response to an individual's atypical bodily characteristics and inabilities resulting from a social environment that is shaped by an unjust lack of attention to people with these bodily characteristics. Gregory argues that the inability theory is not overly individualistic because it does not exclude inabilities that result from an unjust lack of attention. The argument is that the inability theory does consider social injustice, even if only of a certain kind. This clarification, however, does not undermine my previous concern about the fact that Gregory's justification for the inability theory does not consider injustice of the other kind: No reasons are presented that are independent of conservative considerations for excluding inabilities that result from the attitudes of others towards a person's atypical bodily features. Such a more comprehensive justification of condition 2 is also important because there are *pro tanto* reasons to consider the importance of others' attitudes to the experience of disability in a disability account. Against this background, it seems justified to conclude that from the perspective of someone who also accepts the realist criterion, condition 2—in addition to Gregory's definition of the relevant reference classes—is insufficiently justified.

4. Conclusion

The challenges for the inability theory arise from generally plausible expectations for a disability account that are difficult to meet all together. Therefore, the following outline of problems also applies to others who have attempted to formulate a disability account, with the difference that many of them face additional challenges. To develop his proposal into a comprehensive account of disability, Gregory would have to preserve the already mentioned advantages of his proposal while addressing the problems identified above. The modified account should retain the simplicity and intuitive character of the current proposal. It should also make clear why disability is normatively relevant, without defining disability as something bad.

In order to address the shortcomings of the inability theory, we first need more detailed definitions. We need a more precise definition of the range of

typical bodily characteristics and ability levels. We also need a more detailed characterization of the relevant reference classes. Second, as we have seen, the inability theory faces counterexamples that suggest that the theory is overinclusive. One type of such counterexamples concerns intuitively insignificant ability limitations, such as wiggling one's ears. With a little creativity, we can identify associated atypical ability levels for many atypical bodily features. This type of case suggests that statistically atypical bodily features that lead to statistically atypical ability levels are not a sufficient criterion for disability. The suspicion arises that a statistically atypical ability level associated with a statistically atypical bodily feature may be an important indicator of the presence of disability but does not in itself explain what individual cases of disability have in common. Thus, we might conclude that the inability theory has not yet captured the element that could underpin the theory's explanatory power. Third, we need an account of what makes a reference class adequate, without sacrificing the aforementioned advantages of the inability theory. Fourth, either we have to abandon condition 2 and ensure that the theory satisfies the conservative criterion in some other way that is not subject to the same problems, or we have to justify condition 2 on the basis of nonconservative considerations as well. Attempts to meet all these challenges for the inability theory may show us that adopting it is so costly and problematic that this motivates us to look for an alternative account.[31]

Johannes Gutenberg-Universität Mainz
steelsen@uni-mainz.de

REFERENCES

Amundson, Ron. "Disability, Handicap, and the Environment." *Journal of Social Philosophy* 23, no. 1 (1992): 105–18.

Barclay, Linda. *Disability with Dignity: Justice, Human Rights and Equal Status*. Routledge, 2019.

Barnes, Elizabeth. *The Minority Body: A Theory of Disability*. Oxford University Press, 2016.

31 This paper includes ideas and material presented at various occasions in Bern, Southampton, Mainz, and Berlin. The discussions on these occasions were immensely helpful, and I would like to thank more people than I can mention here. I am particularly grateful to Maike Albertzart, Jonas Blatter, Delphine Bracher, Andreas Cassee, Anna Goppel, Alex Gregory, David Heering, Tim Henning, Andreas Müller, Markus Stepanians, and two anonymous reviewers for their invaluable comments and suggestions.

Begon, Jessica. "Disability: A Justice-Based Account." *Philosophical Studies* 178, no. 3 (2020): 935–62.

Buchanan, Allen, Dan W. Brock, Norman Daniels, and Daniel Winkler. *From Chance to Choice: Genetics and Justice*. Cambridge University Press, 2000.

Burch, Susan, and Ian Sutherland. "Who's Not Yet Here? American Disability History." *Radical History Review* 94 (2006): 127–47.

Clare, Eli. "Stolen Bodies, Reclaimed Bodies: Disability and Queerness." *Public Culture* 13, no. 3 (2001): 359–65.

Francis, Leslie, and Anita Silvers. "Perspectives on the Meaning of 'Disability.'" *AMA Journal of Ethics* 18, no. 10 (2016): 1025–33.

Glasser, Adrian. "Presbyopia." In *Encyclopedia of the Eye*, edited by Darlene A. Dartt, Joseph C. Besharse, and Reza Dana. Elsevier Academic Press, 2010.

Gregory, Alex. "Disability as Inability." *Journal of Ethics and Social Philosophy* 18, no. 1 (2020): 23–48.

Grice, Paul. "Logic and Conversation." In *Syntax and Semantics*, vol. 3, *Speech Acts*, edited by Peter Cole and Jerry L. Morgan. Academic Press, 1975.

Harris, Susan S. "Vitamin D and African Americans." *Journal of Nutrition* 136, no. 4 (2006): 1126–29.

Holden Brian A., Nina Tahhan, Monica Jong, David A. Wilson, Timothy R. Fricke, Rupert Bourne, and Serge Resnikoff. "Towards Better Estimates of Uncorrected Presbyopia." *Bulletin of the World Health Organization* 93, no. 10 (2015): 667.

Hull, Richard. "Defining Disability: A Philosophical Approach." *Res Publica* 4, no. 2 (1998): 199–210.

Jones, Patrice, Mark Lucock, Martin Veysey, and Emma Beckett. "The Vitamin D-Folate Hypothesis as an Evolutionary Model for Skin Pigmentation: An Update and Integration of Current Ideas." *Nutrients* 10, no. 5 (2018): 554.

Kahane, Guy, and Julian Savulescu. "The Welfarist Account of Disability." In *Disability and Disadvantage*, edited by Kimberley Brownlee and Adam Cureton. Oxford University Press, 2009.

Kingma, Elselijn. "Naturalist Accounts of Mental Disorder." In *Oxford Handbook of Philosophy and Psychiatry*, edited by K. W. M. Fulford, et al. Oxford University Press, 2013.

———. "What Is It to Be Healthy?" *Analysis* 67, no. 2 (2007): 128–33.

Linton, Simi. *Claiming Disability: Knowledge and Identity*. New York University Press, 1998.

Oliver, Michael. *Understanding Disability: From Theory to Practice*. St. Martin's Press, 1996.

Silvers, Anita. "On the Possibility and Desirability of Constructing a Neutral Conception of Disability." *Theoretical Medicine* 24, no. 6 (2003): 471–87.

Stock-Landis, Leah. "The Toxic Myths I Internalized as a Person with Facial Differences." The Mighty (website), November 10, 2022. https://themighty.com/topic/goldenhar-syndrome/toxic-myths-ableism-facial-differences/.

Stone, Sharon Dale. "Reactions to Invisible Disability: The Experiences of Young Women Survivors of Hemorrhagic Stroke." *Disability and Rehabilitation* 27, no. 6 (2005): 293–304.

Wasserman, David. "Philosophical Issues in the Definition and Social Response to Disability." In *Handbook of Disability Studies*, edited by Gary L. Albrecht, Katherine D. Seelman, and Michael Bury. Sage Publications, 2001.

Wasserman, David, and Sean Aas. "Disability: Definitions and Models." In *Stanford Encyclopedia of Philosophy* (Fall 2023). https://plato.stanford.edu/archives/fall2023/entries/disability/.

World Health Organization. "International Classification of Functioning, Disability and Health (ICF)." Adopted May 22, 2001, last updated 2018. https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health.

———. "International Classification of Impairments, Disabilities, and Handicaps: A Manual of Classification Relating to the Consequences of Disease." Adopted May 1980. https://iris.who.int/bitstream/handle/10665/41003/9241541261_eng.pdf.

Young, Stella. "I'm Not Your Inspiration, Thank You Very Much." TED talk, Sydney, April 2014. https://www.ted.com/talks/stella_young_i_m_not_your_inspiration_thank_you_very_much.

Journal of Ethics and Social Philosophy
VOL. 30, NO. 3 · MAY 2025
https://doi.org/10.26556/jesp.v30i3.3887

PERSPECTIVISM AND RIGHTS

Daniele Bruno

PERSPECTIVISM is the view that what an agent ought to do always needs to be determined relative to this agent's epistemic position. When we have absolutely no way of knowing that a seemingly innocuous action stands to cause some serious harm through an unforeseeable causal fluke, it cannot be that it would be wrong for us to perform this action (though it may be highly regrettable if we do, and others may permissibly stop us in reasonable ways).[1]

Although this view has some important theoretical virtues, it also suffers from what appears to be a crucial flaw: it seems incompatible with the existence of universal claim rights. More specifically, perspectivism turns out to be unable to capture the universality of rights, at least as long as we accept two plausible and widely recognized conceptual claims: *Correlativity* between rights and their corresponding directed obligations; and *Obligation-Reason*, the claim that if *S* has an obligatnion not to ϕ, then *S* has a reason not to ϕ. When combined with plausible statements of universal rights, Correlativity and Obligation-Reason yield claims that necessarily conflict with the basic tenets of perspectivism. Perspectivists thus find themselves in a situation where they must reject either Correlativity, Obligation-Reason, or the existence of universal rights. My first aim in this article is to carefully draw out this *rights trilemma for perspectivism*, showing how it necessarily arises due to structural features of the view.

My second aim is to discuss the options available to perspectivists in reaction to this problem. While the problem of universal rights for perspectivism has been broached elsewhere (albeit often only superficially), in-depth discussions of how perspectivists might be able to deal with it are few and far between until now. The most prominent is due to Michael Zimmerman, who suggests that perspectivists should give up the idea of universal claim rights in its most familiar form.[2] Zimmerman instead proposes that we should focus on a different kind of rights that we may call *perspectival rights*. Whether these rights obtain depends fundamentally on the epistemic position of the obligated

1 Following Michael Zimmerman, perspectivism is also sometimes called *prospectivism* (*Living with Uncertainty* and *Ignorance and Moral Obligation*).

2 See Zimmerman, *Living with Uncertainty*, ch. 2, and *Ignorance and Moral Obligation*, ch. 5. See also Littlejohn, "Is Justification Just in the Head?"

party, which brings rights in line with perspectivist judgments about reasons and oughts. As I shall argue, such a move comes with some very substantial costs, both extensional and theoretical in nature. Careful consideration of cases reveals that judgments about rights simply have a significantly more robust "objective flavor" than judgments about reasons and oughts. Rights thus resist a simple subsumption into the perspectivist framework, giving perspectivists reason to search for alternative solutions to the rights trilemma.

In what follows, I propose such a different approach, unexplored until now. I argue that in response to the rights trilemma, perspectivists should reject Obligation-Reason. This would require us to understand obligation as a *prima facie* normative notion. On this view, rights and obligations generally give rise to reasons but can fail to do so altogether when certain (epistemic) conditions are not met. While rejection of Obligation-Reason may appear to be a radical step, I lay out some considerations that hopefully soften its bite. As I argue, even a *prima facie* view manages to retain two crucial conceptual features of rights—namely, the right holder's entitlement to demand compliance and the existence of important normative consequences in cases of rights infringement (e.g., duties of explanation or compensation). Though the move toward a *prima facie* view of rights and obligations is not without costs, I argue that overall, it represents the best answer to the problem of universal rights for perspectivists, at least when considering the even more costly alternatives.

The article proceeds as follows. In section 1, I offer a slightly more precise definition of perspectivism, employing the notion of *epistemic filters* while also briefly sketching some of the main attractions of the view. In section 2, I lay out the rights trilemma for perspectivism. In section 3, I discuss and criticize Zimmerman's perspectival view of rights, which gives up the idea of truly universal rights. I then develop my alternative response to the rights trilemma in two steps. In section 4, I lay out the *pro tanto* view of rights and show that while it represents an important step toward a solution of the rights trilemma, it cannot itself provide the needed fix. In sections 5 and 6, I then develop and defend the *prima facie* view of rights as a novel way out of the trilemma. I end with a brief outlook in the concluding section.

1. PERSPECTIVISM AND EPISTEMIC FILTERS

1.1. Reasons and Oughts

Any discussion of perspectivism, especially one that focuses on rights and obligations, is well advised to aim for maximal clarity in the conceptual repertoire it draws on. Not only do the terms 'ought', 'reason', and 'obligation' mean different

things to different people, but certain definitional combinations make core perspectivist claims either tautological or downright contradictory. In this first section, I therefore want to lay out a specific understanding of the core perspectivist idea that is put forward in terms of normative reasons and the all-things-considered ought claims that they support. It builds on the widely accepted view of normative reasons as considerations that count in favor of an action or attitude and the idea that these reasons jointly determine what an agent ought to do.[3]

When laying out a perspectivist position in terms of reasons in the way I am going to propose, the notion of reasons itself should be held not to presuppose any controversial features that settle the controversy by conceptual fiat. This means that normative reasons should not be understood as conceptually entailing any kind of mind dependence or subjectivizing element. Holding this conceptually nonsubjective notion of normative reasons fixed, we can then understand perspectivism as making a substantive claim about how the reasons for actions that determine what we ought to do necessarily depend on our perspective.

1.2. Epistemic Filters

On the construal I will rely on in the following, perspectivism claims that there are *epistemic filters* on the normative reasons that determine what we (morally and otherwise) ought to do.[4] To get a grip on what this claim amounts to, it is helpful to consider a classic case of action under ignorance as an illustration of the differences between perspectivism and the contrary view, objectivism.

> *Sugar*: Host has Guest over for tea. Guest politely requests that Host put a tablespoon of sugar in her cup. Unbeknownst to both, the sugar in Host's sugar pot has been laced with an undetectable poison. If Host were to spoon in the sugar into the cup, Guest would die. What should Host do?

3 See paradigmatically Scanlon, *What We Owe to Each Other*, ch. 1. I will leave open the question of how exactly reasons determine requirements. For a thoroughgoing recent proposal, see Schmidt, "The Balancing View of Ought."

4 The term is due to Dancy (*Practical Reality*, 66). For a carefully elaborated version of the filter view, see Kiesewetter, "What Kind of Perspectivism?" I should note here that the question of how to best frame perspectivism is itself contentious. My aim in laying out the filter view is not to take a stance in this intraperspectivist debate. Instead, I simply seek to provide a formalization of the view that is sufficiently clear and suited to bringing out the rights problem in the most efficient manner. That being said, the problems discussed in the following should be straightforwardly translatable to alternative ways of capturing perspectivism, e.g., through the idea of prospective value. See Zimmerman, *Ignorance and Moral Obligation*, ch. 2.

According to objectivism, the question of what Host ought to do must be settled on the basis of all the relevant facts, whether or not they are or could be known by the agent. Given that the objective situation is what matters, an objectivist can confidently conclude that Host ought not to spoon the sugar into Guest's cup. That being said, an objectivist may of course also claim that this does not yet entail anything about Host's *blameworthiness* for doing so. Although acting wrongly, Host may nonetheless be fully excused for her action in light of her blameless ignorance of the pertinent facts.[5]

Perspectivists, on the other hand, disagree with this verdict. For perspectivists, what we ought to do cannot simply be determined on the basis of all facts but instead only on a subset of these—namely, those facts that are epistemically accessible to an agent.[6] Facts that are not epistemically accessible to *S*, or, differently put, facts that *lie outside of S's perspective*, cannot determine what *S* ought to do.[7] Thus, Host ought to give the sugar, since the reasons in favor of doing so (complying with Guest's request, most importantly) outweigh any reasons against that Host is and could be aware of (e.g., that excessive sugar intake is unhealthy). Assuming that what *S* ought to do is exclusively a function of what reasons *S* has, we can capture the core idea behind perspectivism as follows:

> *Perspectivism: X* is a reason for *S* to ϕ only if *X* is epistemically accessible to *S*.

Note that on the filter view, merely apparent *X*s are not suited for playing the role as reasons. Thus, perspectivists of this stripe are not committed to the view that one of Host's normative reasons for acceding to Guest's request is that

5 Traditionally, the sense of ought at issue here has often been identified with the ought of moral obligation. Recently, however, the debate has more commonly been led in terms of a more general kind of ought, the all-things-considered deliberative ought, which conclusively settles the question "What should I do?" I will focus on the latter in what follows. In doing so, I will use the term 'wrong' as meaning *such that* S *ought not to perform* and 'permissible' as meaning *such that it is not that case that* S *ought not to perform*, with 'ought' referring to the mentioned all-things-considered deliberative sense. Though this has the unfortunate effect of suggesting an exclusively moral reading, the stylistic advantages seem to me to outweigh any disadvantages.

6 In this context, perspectivists speak of "available reasons" (Kiesewetter, *The Normativity of Rationality*, ch. 8) or "possessed reasons" (Lord, *The Importance of Being Rational*, 8–9). Since my aim is not to provide a defense of a fleshed-out version of perspectivism, I will leave open the contested question of what it takes for a piece of evidence to be accessible to an epistemic subject.

7 The notion of perspective in play here is quite minimal. An agent's perspective is determined solely based on the evidence epistemically accessible to her. This usage of the term thus differs from the richer and more ambitious concepts discussed in Camp, "Perspectives and Frames in Pursuit of Ultimate Understanding"; and Sliwa, "Making Sense of Things."

the stuff in the sugar jar is pure sugar. After all, even though Host blamelessly believes this to be the case, there is no pure sugar in the jar. Instead, perspectivists following the filter view would identify factual *X*s that are epistemically accessible to Host that can play the role of reasons in this case.[8]

An alternative to this filter view of perspectivism would be a more radical belief-relative perspectivism or subjectivism, which would allow for the normative relevance of merely apparent *X*s. On this view, we would determine an agent's reasons by considering their actual beliefs and then asking ourselves what reasons they would have if their beliefs were true.[9] In what follows, I assume the less extreme (and it appears to me, substantially more plausible) evidence-relative version of perspectivism.[10]

1.3. Some Virtues of Perspectivism

While I will be primarily concerned with a problem for perspectivism in what follows, I would be remiss not to at least briefly mention some of its most important virtues. Otherwise, the project of exploring ways out of the rights impasse for perspectivists might seem moot to begin with—one might simply recommend adoption of objectivism instead. For this reason, I now sketch what I believe to be the most important advantages of perspectivism over objectivism, although the available space permits me only the briefest of glosses of each of them.

First, perspectivism has important extensional advantages over objectivism. Perspectivism can, and objectivism apparently cannot, properly account for the fact that sometimes we ought not to pursue certain courses of action because doing so would be risky or reckless in virtue of our suboptimal epistemic position. Some classic examples for such cases of choices under known uncertainty are Jackson's case of the doctor Jill and Parfit's case of the miners.[11]

A second important advantage of perspectivism over objectivism that is regularly cited is its *action guidingness*. One crucial feature of all-things-considered ought judgments, it may plausibly be claimed, is that they can serve as guides

8 These could be, e.g., summary facts about epistemic probabilities (for all that *S* knew, *p* was true) or individual facts about single bits of evidence, such as the facts that there has always been pure sugar in the sugar jar before, that it was labeled "sugar," and there was no sign of tampering.

9 Such subjective versions of perspectivism are defended by Jackson, "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection"; and Parfit, *On What Matters*, vol. 1, ch. 7.

10 Again, however, the points and arguments will find wider application *mutatis mutandis*.

11 See Jackson, "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection," 462–63; and Parfit, *On What Matters*, 1:159.

to action. Objectivism, which allows for the truth of ought judgments that lie beyond the epistemic ken of the agents subject to them, appears ill equipped to make room for this role of practical judgment.[12]

Third, and finally, Errol Lord argues that only perspectivism can account for the fact that when we ought to do something, it must be at least possible for us to do this thing for the right reasons.[13] Following a similar line of thought, Vuko Andrić argues that while both objectivists and perspectivists can account for the commonly accepted principle of Ought Implies Can, the rationales undergirding this principle, which have to do with the necessity of agents being able to properly react to the right-making features of acts, tell strongly in favor of perspectivism.[14]

2. THE RIGHTS TRILEMMA FOR PERSPECTIVISM

2.1. Universal Rights

Having sketched out the structure and core advantages of the view, let me thus turn to the rights trilemma for perspectivism. To get a grip on the problem, it is necessary to first say a few words on the notion of rights that is at issue. The rights that I am concerned with are claim rights in the classic Hohfeldian sense.[15] What is more, my interest is not with posited rights that are assigned by some political or social body but with the more fundamental and robust moral claim rights.

Some of these rights, although robust in one sense, may nonetheless be substantially conditional. I may have a right against your divulging my saucy secret only if you have promised not to do so or if our past interaction was such that a solid friendship has formed between us. Some moral claim rights may also find application only in specific circumstances (certain kinds of competition, perhaps) or accrue to people with certain roles, such as parents or teachers.

At least some rights, however, appear to be of a more universal nature—these are rights that are held by everyone and that are owed to them by all moral subjects. Plausibly (though not necessarily), these rights may be grounded in certain universal features of the rights holders. We might say, for example, that

12 See, e.g., Way and Whiting, "Perspectivism and the Argument from Guidance"; and Fox, "Revisiting the Argument from Action Guidance."

13 See Lord, "Acting for the Right Reasons, Abilities, and Obligation" and *The Importance of Being Rational*.

14 See Andrić, "Objective Consequentialism and the Rationales of 'Ought' Implies 'Can'" and *From Value to Rightness*.

15 See Hohfeld, *Fundamental Legal Conceptions*.

the most fundamental of these rights accrue to all of us in virtue of our humanity or, perhaps more liberally, our nature as sentient creatures with interests that can be respected or set back. Examples of such rights include rights to freedom of thought and expression, freedom of movement, and freedom from interference with one's privacy and bodily autonomy. Many of these rights have been codified as *human rights*. While the content of our universal rights in the moral sphere may ultimately differ from those enshrined in political declarations such as the UN's Universal Declaration of Human Rights, it is important to note that universality and unconditionality are crucial and irreplaceable elements in both political and philosophical notions of these fundamental rights.[16]

Ultimately, the rights problem for perspectivism that I sketch below arises with respect to every robust form of moral claim right. The arguments that I offer could therefore equally be constructed with examples drawing on more obviously conditional rights, say, rights to the performance of a specific promised act.[17] Since, however, the problem seems especially pressing regarding universal rights, I focus exclusively on these in the following. More concretely, I operate with one specific pet example of a putatively universal right.[18] This example has a relatively well-defined content, which will hopefully not only lend it greater plausibility but also make the construction of clear examples in which to test our intuitions about it easier.

> *Privacy*: If *R* has not consented to *S* reading *R*'s diary, then *R* has a right against *S* that *S* not read *R*'s diary.

I ask any readers who do not take the right in Privacy to be a plausible candidate for a universal right to bear with me for the time being. As will become apparent

16 For discussion of the relation between human rights and moral claim rights, see, e.g., Griffin, "Discrepancies Between the Best Philosophical Account of Human Rights and the International Law of Human Rights"; and Hinsch and Stepanians, "Human Rights as Moral Claim Rights."

17 In fact, given that promises and other normative powers like legitimate orders and consent are commonly defined with respect to their effects on rights and obligations, the arguments below also yield pressure for perspectivists to (at least partly) reconceptualize them to bring them in line with their theory. I leave a more in-depth discussion of this important point as a task for another day.

18 Readers might be surprised to find that the grammatical structure of the right in question is also conditional but should not let this distract them from the intuitive difference between conditional and universal rights just laid out. The consent clause included in Privacy is itself universal to all rights (or almost all, depending on the limits of consent) and thus does not detract from the universality of the right in Privacy. Rights that I have just called substantially conditional, such as the right to having a promise kept, of course would also include such a further condition and thus would be at least doubly conditional in nature.

shortly, the rights trilemma arises as a result of structural features of universal rights and is thus generally independent of their specific content.[19]

2.2. Two Commonly Accepted Features of Rights and Obligations

The rights trilemma arises when we consider rights like that in Privacy in conjunction with two very plausible conceptual claims. The first of these is:

> *Correlativity*: *R* has a right against *S* that *S* not ϕ if and only if *S* has an obligation not to ϕ owed to *R*.

Correlativity follows straightforwardly from the Hohfeldian conception that is at heart of the modern understanding of claim rights. The nature of the rights at issue is that they provide their bearers with what we may call a form of normative protection. Rights constitute a sort of *normative shield* against others—they rule out certain ways others could act against rights holders as incompatible with respecting the rights.[20] However, rights can be said to do so only if they correspond to normative restrictions on the side of these others.

It thus must be the case that anyone acting contrary to a right finds themselves subject to an obligation not to do so. What is more, this obligation must be directed. After all, the obligated party owes it to the right holder specifically not to perform the act against which the right protects. If I impermissibly read your diary, I not only act wrongly, but I wrong you. Though Correlativity may in principle be challenged, this would require a thorough and extremely revisionary reconceptualization of rights. For this reason, I do not pursue arguments against it in what follows.[21]

The second conceptual claim regards the obligations that correspond to rights. I have said that these constrain the permissible options available to those that are subject to these obligations. In what way do they constrain them? One simple and straightforward way one might take this to happen is the following.

> *Obligation-Ought*: If *S* has an obligation not to ϕ, then *S* ought not to ϕ.

19 Readers who have doubts about the content about Privacy on the grounds that its current form makes it vulnerable to counterexamples involving countervailing considerations like emergencies are referred to the discussion in section 4 below. There, I will turn to the question of whether Privacy needs to be specified further to be extensionally adequate.

20 For the shield metaphor, see, e.g., Regan, *Animal Rights, Human Wrongs*, 26; and Schauer, "A Comment on the Structure of Rights," 229–31.

21 For a defense of Correlativity in a similar dialectical context, see Zimmerman, *Ignorance and Moral Obligation*, 119–23.

Obligation-Ought is a claim that has been explicitly defended by several prominent moral philosophers.[22] Nonetheless, it is the more controversial of the two claims provided here. As I later show, the rights trilemma can also be constructed with a weaker claim regarding the normative force of obligations. This is the claim that I call Obligation-Reason.

> *Obligation-Reason*: If *R* has an obligation not to ϕ, then *R* has a (normative) reason not to ϕ.

However, I shall first proceed on the stronger claim of Obligation-Ought. I do this for the simple strategic reason that it allows me to reconstruct the rights trilemma in the clearest and most transparent way possible. We return to the possibility of rejecting Obligation-Ought and employing Obligation-Reason as an alternative in section 4 below.

2.3. The Rights Trilemma for Perspectivism

Perspectivism cannot allow for the joint truth of Privacy, Correlativity, and Obligation-Ought. Perspectivists thus are forced to jettison one of these three intuitively highly plausible claims. This is the rights trilemma for perspectivism. To see how the inconsistency arises, consider first that the three propositions just outlined together logically entail the following:

> *Ought Implication*: If *R* has not consented to *S* reading *R*'s diary, then *S* ought not to read *R*'s diary.

That the Ought Implication conflicts with perspectivism is quickly shown. Take the following restriction, which follows straightforwardly from the basic account of perspectivism laid out above.

> *Perspectivist Restriction:* If *S* ought not to ϕ on account of *p*, then *p* must be epistemically accessible to *S*.

We now need only to take on board the possibility of cases of ignorance regarding the application conditions of a given universal right. For Privacy, this would be:

> *Ignorance*: It is possible that *R* has not consented to *S* reading *R*'s diary, yet *S* has no epistemic access to this fact.

22 For some prominent defenses, see, e.g., Shafer-Landau, "Specifying Absolute Rights"; and Wallace, *The Moral Nexus*. At least if we assume the truth of a kind of moral rationalism and of Correlativity, this view is also implied by some of the most prominent theories of rights, e.g., the view of rights as side constraints. See, e.g., Nozick, *Anarchy, State, and Utopia*; and Dworkin's view of rights as trumps in *Taking Rights Seriously*.

The truth of Ignorance is quickly established by means of example.

> *Twin Impersonation*: Sam has long had an interest in reading his flatmate Rayan's diary. However, Rayan has never given Sam consent to do so. One day, Rayan's identical twin sister Riya, of whose existence Sam had no clue, pretends to be Rayan and approaches Sam, telling him she is fine with him having a look and informing him of the diary's location in the drawer of Rayan's bedside table.

Perspectivist Restriction and Ignorance in turn logically entail the following:

> *Perspectivist Implication*: It is possible that *R* has not consented to *S* reading *R*'s diary without it being the case that *S* ought not read *R*'s diary.

Perspectivist Implication and Ought Implication stand in clear and undeniable contradiction to each other. Since both the truth of Ignorance and the entailment relations laid out cannot reasonably denied, perspectivists must reject one of Privacy, Correlativity, and Obligation-Ought—a trilemma. Given the plausibility of each of the three claims that make up the horns, this is a serious drawback for the view. Just how serious a drawback it is, however, depends on how high the costs of rejection are for each of these claims. I now turn to this question.

3. A PERSPECTIVAL VERSION OF RIGHTS?

3.1. Rejecting Privacy as a Solution to the Rights Trilemma

As mentioned before, the most prominent extant proposal on how perspectivists are to deal with something like the rights trilemma is due to Michael Zimmerman.[23] Zimmerman opts for a wholesale rejection of universal rights, which solves the rights trilemma by denying the truth of Privacy. Zimmerman is of course aware of the importance that rights have in our lives and thus does not propose that we give up talk and thought about rights altogether. Instead, he suggests that perspectivists reconceptualize the category from the ground up, bringing it in line with the basic tenets of their view. On this view, which we may call *rights perspectivalism*, people are strictly speaking not bearers of rights against violations of their privacy, bodily integrity, freedom of speech, etc., but instead can accurately be said to have only rights not to be subjected to *epistemic risks* of such violations.[24] To illustrate this, consider this perspectival version of Privacy:

23 See Zimmerman, *Living with Uncertainty*, ch. 2 and *Ignorance and Moral Obligation*, ch. 5.

24 Zimmerman's discussion is framed by a general statement of rights not to be harmed, the Harm Thesis, which he rejects. Instead, he defends the Risk Thesis, which implies

> *Privacy*$_{Persp}$: If, *according to S's perspective*, *R* may not have consented to *S* reading *R*'s diary, then *R* has a right against *S* that *S* not read *R*'s diary.

Unlike its universalist cousin, Privacy$_{Persp}$ does not lead to a contradiction with Perspectivist Implication when conjoined with Correlativity and Obligation-Ought. And as Zimmerman points out, the general idea that we have rights not to be exposed to (even merely epistemic) risks of harm is hardly outlandish. We plausibly have a right not to be served highly perishable food that the chef knows has not been refrigerated for a prolonged period even if, by chance, no dangerous bacteria have yet formed on it. We can formally state the position as follows:

> *Rights Perspectivalism*: *R* has a right against *S* that *S* not ϕ only if all factors relevant to this right's obtaining are epistemically accessible to *S*.

On rights perspectivalism, rights are therefore also subject to an epistemic filter, just as reasons are on perspectivism. This allows the view to avoid the potential mismatch between rights and oughts that gives rise to the rights trilemma. However, it manages to do so only at very high theoretical cost. In what follows, I highlight three pressing problems for rights perspectivalism: two extensional shortcomings and one more fundamental theoretical issue concerning the complexity of rights.

3.2. Undergeneration Worries

A first and perhaps most obvious worry is the perspectival view's *undergeneration* of rights. On rights perspectivalism, individuals lack rights in many circumstances in which we would intuitively take them to hold them. This class of cases evidently includes Twin Impersonation, as it must to avoid the problems described in the last section. This already puts the perspectival view out of line with most people's considered intuitions. Surely, it is not the case that Rayan's right to privacy protects her only against intrusive peeks into her diary by Sam when her twin refrains from interfering yet stays silent as soon as the twin's actions have affected Sam's epistemic situation. Rights, it seems, are supposed to be more robust than that.

Quite generally, one might take issue with the way in which the perspectival view allows for a loss of our rights protection through events over which we have no power and of which we may have no knowledge. The problem is not that rights fail to protect us under all circumstances—I can waive my rights through consent or plausibly also forfeit them through some impermissible actions. On the perspectival view, however, the normative protection provided

---

perspectival rights in the sense laid out below.

by rights will not only obtain in far fewer cases but regularly fail in situations where the failure is completely independent from my own actions.

Even more serious worries of undergeneration arise if we allow for the possibility of agents' perspectives yielding false normative claims, as Zimmerman in fact allows.[25] If rights protect us only against acts that appear wrong from a would-be perpetrator's perspective and if perspectives on normative facts can be misleading, then we face the possibility of widespread rights erasure in normatively wayward societies. For one thing, this means that chattel slaves in the ancient Hittite kingdom did not have a right to freedom if the slaveholders' culturally ingrained ignorance of the moral impermissibility of slavery was so deep that the correct moral verdict was not accessible to the slaveholders.[26]

Even worse, it also has implications for the beliefs of the Hittite slaves themselves. Imagine that some of these slaves started to deliberate about the morality of their own situation, setting off from perfectly true premises about factors like their human dignity and the intrinsic value of autonomy. Imagine then that these slaves finally came to the conclusion that they had a moral right to be free from the yoke of slavery. Ironically, rights perspectivalists would have to claim that these slaves would be mistaken in this conclusion. After all, it is a *slaveholder's* epistemic position that determines which rights a slave possesses, since the former would be the potential holder of the obligations corresponding to the latter's putative right to freedom. On a perspectival view of rights, a liberation movement started by such slaves and aimed at a change of public opinion in Hittite society could therefore not accurately be described as having the aim to have their preexisting moral rights respected by the Hittite slaveowners. Instead, it would have to be understood as a movement aimed at *bringing into existence* the very moral rights that intuitively ground the righteousness of their project in the first place. In addition to the intrinsic objectionability of denying the slaves a right to freedom, rights perspectivalism seems to get things backwards in a seriously unsatisfactory manner with respect to this.[27]

3.3. Overgeneration Worries

Besides the mentioned worries about undergeneration, rights perspectivalism also *overgenerates* rights in objectionable ways. On the perspectival view of

25 See Zimmerman, *Ignorance and Moral Obligation*, 63–65.

26 The familiar example of ancient Hittite slaveholders stems from Rosen, "Culpability and Ignorance," 64–65.

27 Note that the objectionability rides exclusively on the perspectivalist idea that these slaves would have to bring about a change in their *moral* rights. A liberation movement will of course correctly aim at a change to the slaves' legal and political rights but is precisely justified because doing so will bring them in line with preexisting, independent moral rights.

rights, we can have rights against others to perform actions that we know to be extremely undesirable for us. Let me illustrate this possibility with an example:

> *Penicillin*: Fang is on holiday in a remote area of a foreign country and contracts a serious infection that needs immediate treatment. Fang goes to see a doctor and tries to let the doctor know that she suffers from a fatal allergy to penicillin and therefore by no means must be treated with this drug. However, due to a language barrier, Fang does not get her point across. The doctor, in line with the best evidence available to her, takes Fang's signals of distress to simply be a plea for urgent treatment and administers penicillin to treat the infection.

Let us assume, uncontroversially, that Fang has a right to good treatment by the doctor. Since on rights perspectivalism, rights are relativized to the epistemic position of the obligated party and not of the claim holder, defenders of this view are forced to conclude that in the situation just described, Fang has a right against the doctor that she be administered penicillin. This is true despite Fang having every reason to hope that the doctor does not administer penicillin. What is more, Fang's appropriate preference against being administered penicillin is grounded in the *very same* interest in health that plausibly undergirds the right to good treatment. Again, the verdict returned by rights perspectivalism turns out to be strongly counterintuitive.

Note that the issue here is not simply that the doctor is morally required to give the penicillin in such a situation. A judgment to this effect alone, perspectivists might and indeed should argue, would be perfectly in keeping with her duties as a medical professional. Instead, the issue lies with the specific notion of a *right* to treatment with penicillin. The very point of rights is that they can be sensibly asserted by their holders, and that those subject to the rights can be demanded to comply with them. The right to be given penicillin, however, does not appear to be of a kind that Fang could sensibly assert and demand compliance with *as a right to good treatment*, since it is completely free-floating from (and even contrary to) her actual health-related interests.[28] It is therefore a mistake to interpret the right to good treatment in the perspectival fashion, at least in the case of Penicillin.

3.4. Two Kinds of Rights

Although we should not interpret the right to privacy and the right to good treatment in a perspectival fashion, that does not mean that there are no rights

28 There might be other rights at issue that Fang could sensibly assert. I turn to these in the following subsection.

that depend on the obligated party's epistemic situation. As I will argue now, we plausibly have both a right to privacy and a right not to be exposed to (epistemic) risks of privacy violation. This means that $\text{Privacy}_{\text{Persp}}$ and Privacy are both genuine, but they are separate rights, protecting their rights holders against different injuries. While the point of Privacy is to protect individuals from certain outcomes (i.e., from actual breaches of privacy), $\text{Privacy}_{\text{Persp}}$ guards rights holders against a kind of disrespect: by subjecting *R* to an (epistemic) risk, *S* expresses a kind of disrespect to *R* that represents a (lesser) injury even if this risk does not materialize. In acting in a way that *S* has reason to believe will violate *R*'s privacy, *S* expresses an objectionable form of disregard to *R*'s moral standing—an expression that *R* might wish to guard themselves against.

We can see the difference between these two separate kinds of rights clearly when we consider how they come apart in various situations. We have already encountered an example of an action infringing on Privacy without infringing on $\text{Privacy}_{\text{Persp}}$ in the case of Twin Impersonation. On the other hand, we may imagine a case in which *R* has in fact validly consented, but *S* is unsure whether valid consent was given (perhaps *S* has a suspicion that *R* was under duress while giving consent). In this situation, *R* might still rightfully feel wronged by *S* if *S* reads her diary. After all, *S*'s action expresses an objectionable disregard of *R*'s interest, even though *R*'s privacy is not actually violated. Finally, in the worst but probably most common case, *S* might violate both Privacy and $\text{Privacy}_{\text{Persp}}$, for example if *S* consciously defies *R*'s wishes in full knowledge of the circumstantial factors.

Importantly, the differences between the various cases just mentioned are reflected in different kinds of redress that are appropriate. When perspectival rights are violated, the wrongdoer's duties of redress plausibly focus on a sincere apology that underscores the appreciation for the moral standing of the rights holder, since the main issue with the action performed was the fact that it expressed a serious disrespect for this moral standing. Redress for infringements of nonperspectival rights, on the other hand, involve a focus on compensation for the actual harm and damages that are brought about by the action, while violations of both rights combined require the most stringent forms of redress that involve both elements.

Where rights perspectivalism goes wrong is therefore not in positing perspectival rights but rather in denying nonperspectival rights. In identifying these two kinds of rights with each other and opting for the reality of only the former, rights perspectivalists efface the intuitive difference between the two. Here, I have of course only barely grazed the rich and complex structure of the various kinds of rights we have and the ways in which others can fail to fully respect them. What I hope has become apparent, however, is that we cannot

do justice to this richness and complexity if we reduce all rights to only one kind of right, as rights perspectivalism demands of us. This gives us a further reason to reject the view.

3.5. The Costs of Revisionism

In sum, the perspectival view of rights saddles us with broad-scale revisionism about both the nature and extension of rights, and this is a serious cost to the view. Still, it is important to acknowledge that despite the intuitive implausibility of the results reached, this does not yet constitute a knockdown argument against perspectivalism. It may still be that, in sum, the costs just outlined are ones we must bear, given the downsides to the alternatives. These alternatives are the other two horns of the trilemma, on the one hand, and the options of rejecting perspectivism for objectivism, on the other. With that, I now turn to what I claim is ultimately the best option for perspectivists—namely, rejecting Obligation-Ought and its weaker alternative, Obligation-Reason.

4. THE *PRO TANTO* VIEW OF RIGHTS AND OBLIGATIONS

Let me begin with Obligation-Ought. Above, I already acknowledged that this claim is more controversial than might appear at first glance. In this section, I briefly lay out and defend an alternative view, which we may call the *pro tanto view of rights and obligations*. As we shall see, rejecting Obligation-Ought in favor of this view, although by no means an unattractive move, will not by itself solve the rights trilemma. The *pro tanto* view of obligations turns out to be compatible with the weaker claim of Obligation-Reason, which still suffices to construct a rights trilemma. However, laying out and defending the *pro tanto* view of rights still proves to be of value, since it provides the basis for developing the *prima facie* view of rights and obligations as a successful solution to the trilemma.

The *pro tanto* view of rights has its most prominent defenders in H. L. A. Hart, Joel Feinberg, and Judith Jarvis Thomson.[29] Its basic idea is well expressed by Bernard Williams:

> We should recall that what is ordinarily called an obligation does not necessarily have to win in a conflict of moral considerations.[30]

29 See Hart, "Are There Any Natural Rights?"; Feinberg, "Supererogation and Rules"; and Thomson, *The Realm of Rights*. Some more recent defenders of the *pro tanto* view include Frederick, "*Pro Tanto* Versus Absolute Rights"; Rettig, "Rights and Practical Reasoning"; and Kiesewetter, "*Pro Tanto* Rights and the Duty to Save the Greater Number."

30 Williams, *Ethics and the Limits of Philosophy*, 180.

As the name suggests, the *pro tanto* view takes both 'right' and the corresponding term 'obligation' to express *contributory notions*. As such, obligations can compete against other considerations in determining what an agent ought to do, all things considered, in much the same way that ordinary moral reasons compete.[31] In this sense, the *pro tanto* view allows for truly universal claim rights, since their status as considerations with contributory force is independent from many contextual features, in particular from the question of what could speak for alternative courses of actions. On the *pro tanto* view, rights and their corresponding obligations *always* have a weight, though they may sometimes be outweighed by other rights, and perhaps even non-rights-related considerations.

Its ability to make sense of the existence of widespread conflicts of obligations is one of the most important reasons to reject Obligation-Ought and embrace the *pro tanto* view of rights. To illustrate such a conflict, take the following case.

> *Blood Type*: Sam has long had an interest in reading his flatmate Rayan's diary. However, Rayan has never given Sam consent to do so. One day, Rayan's twin sister Riya suffers a serious injury and needs urgent medical treatment. The paramedics call Rayan's flat and ask Sam about the by now unconscious Riya's blood group. Sam knows that Rayan and Riya have the same blood group and that information on Rayan's blood group could be found somewhere in Rayan's diary, which he knows is in her bedside table.

I think that intuitions are almost univocally clear that in Blood Type, Sam is permitted to read Rayan's diary. On the *pro tanto* view, this can be explained in the following way: although Sam has an obligation toward Rayan not to read her diary, he has a weightier obligation toward Riya to do what is necessary to save her life. In Thomson's words, the obligation toward Rayan to respect her right to privacy would therefore be *permissibly infringed* if Sam were to access her diary in this situation.

This way of thinking about rights may give rise to a worry. In going *pro tanto*, are we not giving up a core feature of rights—namely, the special normative force that they are supposed to have? Surely, the obligations corresponding to moral rights are not just garden-variety moral reasons like any other but particularly important moral considerations that have what we may call *peremptory force* or *special stringency*.[32]

31 That obligations and ordinary moral reasons compete in the same way does not mean that an obligation to ϕ just is a simple moral reason to ϕ. I will return to this point presently.

32 For peremptory force, see Raz, *The Morality of Freedom*. For special stringency, see Owens, *Shaping the Normative Landscape*. I thank an anonymous referee for JESP for pressing me to explicitly address this issue.

It is true that in rejecting Obligation-Ought, the *pro tanto* view of rights gives up the most straightforward and indisputable way of establishing a special normative force for rights and obligations. However, this does not mean that in adopting this view, we lose the ability to distinguish obligations from other moral reasons. Even less does it commit us to the claim that the former will be regularly outweighed by the latter. The *pro tanto* view is committed only to rights *sometimes* being permissibly infringed. It leaves open how frequently such infringements will occur and what type of normative considerations can outcompete obligations. For all that is implied by the view, rights may very well regularly override most other normative considerations, being infringed only in relatively rare circumstances and perhaps even only on account of other moral rights.[33]

There are several ways in which to theoretically capture the peremptory force of obligations within a nonabsolutist framework. Perhaps the most prominent and promising one is Joseph Raz's notion of an *exclusionary reason*.[34] For Raz, an exclusionary reason is a "reason to refrain from acting for some reason," and as such, it also excludes the normative strength of the reasons that it forbids from consideration.[35] Still, exclusionary reasons must be weighed against non-excluded reasons should these come into conflict. Understanding obligations as exclusionary reasons along Razian lines gives us the theoretical resources to accurately account for most of the ways that rights intuitively constrain us, without committing to something like Obligation-Ought. While much more would admittedly need to be done to flesh out the details of this view, rights and obligations thus can retain a plausible kind of peremptory force even on the *pro tanto* view.[36]

4.1. Specificationism and the Infringement View

What would be the alternative to the infringement view involving *pro tanto* rights in accounting for cases like Blood Type? If we want to uphold something along the lines of a general right to privacy but hold onto both Obligation-Ought

33 There is a further important element to rights and obligations that sets them apart from other moral reasons: they leave a *moral residue* even when outweighed, giving rise to claims to explanation and compensation. I discuss this further below.

34 One pertinent alternative way of explaining the peremptory force of obligations without subscribing to Obligation-Ought is David Owens's habit-based account (*Shaping the Normative Landscape*, ch. 3).

35 Raz, *Practical Reasons and Norms*, 39.

36 For a fully fleshed-out account of the role of *pro tanto* rights in deliberation, drawing centrally on the notion of exclusionary reasons, see Rettig and Fornaroli, "Conflicts of Rights and Action-Guidingness."

and the intuition that Sam is permitted to read the diary in Blood Type, our only option appears to be to specify the right to privacy in such a way that it does not cover Blood Type to begin with.[37] Such an all-things-considered right to privacy would have to look somewhat like the following:

> *Privacy*$_{ATC}$*:* If *R* has not consented to *S* reading *R*'s diary and reading the diary saves no lives that require saving and does not prevent serious international diplomatic incidents and . . . , then *R* has a right against *S* that *S* not read *R*'s diary.

The list of further necessary restrictions included in the placeholder ellipsis is as long as will be necessary to cover the range of possible situations in which we wish to uphold an intuitive judgment that reading a diary without consent would be all-things-considered permissible. On such a *specificationist* view of rights, the problem with cases like Blood Type is solved by simply denying that there is any conflict of rights.[38] In fact, on specificationism, rights and obligations *never* come into conflict, at least not in the sense of individual obligations yielding overlapping and contradictory recommendations in any given situation. Rather, a large number of precisely specified individual rights form a vast interlocking whole without overlaps, much like a perfect jigsaw puzzle.[39] For this reason, we will always be able to isolate a singular right (and a corresponding obligation) that is fully applicable and explains the impermissibility of a certain right-violating course of action.

Specificationism is subject to several problems. For reasons of space, I can only offer a relatively brief gloss of what appear to me three particularly pressing issues. While I regret not being able to do full justice to all possible ways for specificationists to respond to the charges I level, I hope that an overview of the

37 An alternative way of denying the *pro tanto* view—which, however, does away with the idea of a general right to privacy—would be an absolutist form of particularism. On such a view, truths about rights must always be understood relative to particular circumstances, so that there can only ever be a right not to be treated in this precise way in this specific situation. Such rights need not contain counterfactual exception clauses but also cannot be outweighed by other considerations. See Shafer-Landau, "Specifying Absolute Rights," 213. I do not further pursue this possibility here because it not only does away with the idea of universal rights that motivates much of the discussion in this article but also is subject to the same objections I field against specificationism below. I thank Benjamin Kiesewetter for pressing me to consider this possibility.

38 See, e.g., Shafer-Landau, "Specifying Absolute Rights"; Wellman, "On Conflicts Between Rights"; Oberdiek, "Lost in Moral Space" and "Specifying Rights Out of Necessity."

39 For the jigsaw puzzle image, see Wenar, "Rights," sec. 5.2.

most pertinent problems of specificationism at least goes to show that adoption of the *pro tanto* view is a sufficiently attractive option for perspectivists.[40]

A first point of criticism simply regards the implausibly complex nature of rights under specificationist assumptions. Since any right properly conceived would have to include a sufficiently large number of fine-grain specification clauses, specificationists are faced with the charge that no philosopher, let alone any layperson, could accurately claim themselves to know even one simple right that they possess. While specificationists are of course still free to acknowledge that our common "loose talk" of rights has its uses and need not be abandoned, this implication of the in-principle unknowability of our rights is a serious cost to the view.[41]

A second important point regards what defenders of the *pro tanto* view often call the "residue" of permissibly infringed rights.[42] Take Blood Type again. If Rayan comes back later in the afternoon and finds that Sam has rifled through her diary, it seems that the events are not morally neutral with respect to the relationship between the two. At the very least, Sam owes Rayan a lengthy explanation, perhaps even an apology. In other cases of permissibly infringed rights, like Feinberg's famous case of the hypothermic wanderer forcefully breaking into your warm but firmly locked cabin to save his own life, compensation may furthermore be owed for fully permissible infringements of rights.[43] The *pro tanto* view has an easy way of accounting for this residue—it

40 For more fully fleshed-out elaboration, see Thomson, *The Realm of Rights*, 82–104; and Frederick, "*Pro Tanto* Versus Absolute Rights." For a concise summary, see also Wenar, "Rights," sec. 5.2.

41 Some specificationists deny that the number of exception clauses their view must countenance is excessively large. Russ Shafer-Landau claims that "the specificationist might instead insist that the relevant exceptive clauses be relatively few in number and couched in terms of repeatably instantiable kinds of exceptions" ("Specifying Absolute Rights," 212). This suggestion is subject to a dilemma, however. Consider a general repeatably instantiable exception clause to a general right to privacy that forbids violation *unless a significant harm can be avoided*. Either the level of harm necessary is specified in absolute terms, or it is specified in terms relative to the importance of privacy infringement. If it is specified in absolute terms, extensionally implausible results follow. We simply will not be able to find an appropriate level of harm such that it or any harm above it *always* justifies any kind of privacy violation, while a lesser level of harm can *never* justify any infringement of privacy. If the level of harm necessary is specified in terms relative to the importance of privacy infringement, however, we reintroduce a weight to the right to privacy against which the importance of the harm must be compared. This reveals that we have given up the absolute conception of rights inherent to specificationism and landed with a framing that is much more amenable to the *pro tanto* view of rights.

42 The locus classicus is Thomson, *The Realm of Rights*, 84–87.

43 See Feinberg, *Rights, Justice, and the Bounds of Liberty*, 230.

is explained by the rights to privacy and property, which, though outweighed by competing considerations, remain in place and normatively relevant in the described situations.

Specificationism, according to which Blood Type and Feinberg's example of the cabin break-in are not situations in which any right to privacy or right to property find application, has a significantly harder time fully explaining the intuitive residue in these cases. While specificationists are free to invoke independent rights to compensation and explanation to reach extensionally correct descriptions of the situation after the break-in into the cabin, doing so in a satisfactory manner is a much more challenging task than it is for advocates of the *pro tanto* view.

For illustration, take John Oberdiek's prominent proposal, according to which specificationists should account for moral residue by drawing on the more general phenomenon of value pluralism.[44] Oberdiek points to the fact that in situations such as Blood Type and Feinberg's cabin, different values are at stake in the respective incompatible options—health and privacy, and health and property, for example. Sam and the hiker thus find themselves in situations where they must choose between values. They cannot avoid causing some damage to something valuable, even if they choose perfectly. This fact, Oberdiek claims, can account for moral residue in these situations without giving up the specificationist idea that the agents acted in full accordance with all their obligations.

However, Oberdiek's proposal overgeneralizes, for the dynamic described by him does not only occur in situations in which rights are at stake. Consider a person who has decided to donate half of their moderate income to charity. To do so, they must choose between several highly deserving charities. They could support health care campaigns in developing countries, contribute to fighting anthropogenic climate change, or provide financial support to victims of war and persecution. In assigning their donation, this person likewise finds themselves forced to choose one value at the expense of others. This fact is of course not morally neutral and, as such, leaves some sort of moral residue in the form of reason for regret. However, given the supererogatory nature of the would-be donor's intent, this moral residue is substantially different from that faced in Blood Type and Feinberg's cabin. The latter cases give rise to a much more significant kind of moral residue that calls for not only regret but explanation, compensation, and potentially more.

44 See Oberdiek, "Lost in Moral Space," 331–34. Other attempts to explain away intuitions of residue, such as the utilitarian justification that Russ Shafer-Landau adduces ("Specifying Absolute Rights," 216–17), face even more obvious extensional worries, which for reasons of space I cannot go into here.

The mere appeal to value pluralism that Oberdiek provides cannot do the necessary explanatory work to account for the important differences between these two types of cases. The *pro tanto* view, on the other hand, affords a straightforward explanation that draws on the intuitively most pertinent difference: Rayan has a right against Sam not to have her diary read, while the individual charities lack a right against the would-be donor to receive the donation. Of course, specificationists could in principle offer a further account that supplements or altogether supplants the value pluralist proposal. However, the dialectical onus is squarely on the shoulders of specificationists here, as they must find a proposal that matches the intuitive and straightforward explanation offered by the *pro tanto* view.[45]

What is more, there are more general reasons to doubt that this burden can be met. Any account drawing on an alternative source for the explanation of moral residue in cases such as Feinberg's cabin is likely to struggle in capturing the intuitively tight explanatory connection between the putative rights infringement and the obligation to explain and/or compensate. This becomes obvious when we consider a case in which the hiker immediately compensates the owner after the break-in.[46] On the specificationist picture, all rights are fully respected here, and any alternative source of a duty to compensate is satisfied—we seem to be faced with a morally neutral situation. However, this seriously misdescribes the moral dynamic as it intuitively presents itself. It is precisely because some form of (overall permissible) moral infraction has occurred that compensation and explanation are owed. Only the *pro tanto* view can properly account for this tight explanatory connection and thus explain why even immediate and full compensation does not leave us with a morally neutral situation.

Third and finally, specificationism cannot do justice to the way that rights feature in moral deliberation. On specificationism, obligations are something like the output of moral deliberation: it is only once we have fully considered the intricacies of the specific situation that we find ourselves in and have thus ruled out all potential exclusion clauses that we can confidently conclude for a certain right to be pertinent to the decision we are facing. For example, if Sam wants to determine whether Rayan has a fully specified right against him not to read her diary in Blood Type, he would first have to determine whether or not it would be permissible for him to read her diary given the emergency faced by Riya. But once this latter task is done, practical deliberation is already successfully concluded. No conceptual space remains for consideration of any

45 For a much more extensive defense of the moral residue objection to specificationism, see Botterell, "In Defence of Infringement."

46 For this point, see also Frederick, "*Pro Tanto* Versus Absolute Rights," 392.

right to contribute in a substantial way to the success of Sam's reasoning. Rights thus appear explanatorily idle.[47]

In response, specificationists could argue that the role of rights in practical deliberation is different. Hypotheses about rights might instead function as goal-setting mechanisms of practical reasoning, drawing our attention to reason-giving considerations we might otherwise have missed.[48] However, this still renders rights largely dispensable in our practical reasoning. If the point of rights is simply, as Wellman claims, that a given right "marks, rather than explains, the relevant moral reasons," it is not clear why this function could not be equally well fulfilled by other considerations, such as directly considering the pertinent interests, relationships, etc.[49]

It barely bears mentioning that in stark contrast to this assumption, common thought would have rights occupy a very different role. Rights are usually taken to be elements that actively feature in our practical deliberation—they serve as premises in practical reasoning and play both justificatory and explanatory roles for our practical conclusions. When faced with situations of moral conflict such as Blood Type, it is natural to weigh rights against each other. We want Sam to take seriously both Rayan's right to privacy and Riya's right to rescue, and we want a decision based on the proper consideration of both.[50] Likewise, an advisor might simply point to rights that to them appear pertinent in a situation (perhaps reminding Sam of a promise he made toward Rayan in the past), without thereby prejudging the result of Sam's practical reasoning.

This kind of familiar consideration of moral rights within practical deliberation is also important in a number of political and legal contexts. Rights serve crucial explanatory and justificatory roles in establishing important normative conclusions in these fields.[51] Far from being idle wheels, they are the building blocks upon which many an influential argument is constructed. In the role it

47 The canonical formulation of the charge of explanatory idleness is found in Thomson, "Self-Defense and Rights." Rettig provides a well-developed and convincing elaboration of the charge that specificationism deprives rights of a significant role in practical reasoning ("Rights and Practical Reasoning").

48 See Wellman, "On Conflicts Between Rights," 281–82.

49 Wellman, "On Conflicts Between Rights," 282.

50 How exactly this is to be achieved is of course itself a difficult question. For a recent attempt at providing a method by which to guide reasoning in addressing conflicts of rights, see Rettig and Fornaroli, "Conflicts of Rights and Action-Guidingness."

51 We have seen this above, discussing the Hittite slaves in the context of undergeneration worries for rights perspectivalism.

assigns to rights in practical reasoning, specificationism thus seems to simply have things backwards, giving us further reason to reject it.[52]

4.2. Why Rejecting Obligation-Ought Is Not Enough

We thus have good reason to reject Obligation-Ought, one of the three horns of the trilemma as originally constructed. However, as mentioned, this is not enough to save perspectivism from the problem of universal rights. The reason for this is that the rights trilemma can be reconstructed without Obligation-Ought. Instead, we need only the weaker Obligation-Reason:

> *Obligation-Reason*: If *R* has an obligation not to ϕ, then *R* has a reason not to ϕ.

Obligation-Reason is a claim that is not only consistent with the *pro tanto* view of rights but even plausibly entailed by it, at least assuming that all truly normative *pro tanto* notions bottom out in or at least essentially involve reasons. To construct the rights trilemma on the basis of Obligation-Reason, a further slight modification is needed:

> *Perspectivist Restriction*$_2$: If *p* is a reason for *S* not to ϕ, then *p* must be epistemically accessible to *S*.

Although Perspectivist Restriction$_2$ is not strictly implied by every version of perspectivism, I believe that it is nonetheless hard to deny for any plausible version of the view. After all, it is only slightly more specific than the original restriction.[53] As long as perspectivism claims that oughts are perspectival because there is an epistemic filter on its constituents, and we furthermore assume that what we ought to do is determined by contributory considerations that either involve or bottom out in reasons, then perspectivism is firmly committed to Perspectivist Restriction$_2$. In fact, one of the most prominent defenders of perspectivism in recent years, Benjamin Kiesewetter, explicitly defends a view that clearly entails Perspectivist Restriction$_2$.[54]

Perspectivists who want a way out of the rights trilemma that does not involve giving up universal claim rights thus must jettison not only Obligation-Ought but also the much weaker and more plausible Obligation-Reason.

52 I thank an anonymous referee for *JESP* for pressing me to address the challenge that specificationism presents for my arguments in this article in much greater detail than I did at first.

53 However, perspectivists might avoid commitment to Perspectivist Restriction$_2$ by introducing further distinctions (for example between possessed and unpossessed or objective and subjective reasons). I consider this possibility in section 6 below.

54 See Kiesewetter, "What Kind of Perspectivism?"

Though it might seem like a nonstarter at first, I believe this in fact turns out to be a live option—and indeed the best one available to perspectivists. In the final two sections, I lay out my reasons for this belief, first simply sketching the structure of the view and then assessing its merits.

5. THE *PRIMA FACIE* VIEW OF RIGHTS AND OBLIGATIONS

5.1. Extending Rights Infringments

Let us take a step back and reassess the problem. The rights trilemma for perspectivists arises due to the possibility of cases of ignorance such as Twin Impersonation. To successfully uphold perspectivism and some version of a right to privacy, perspectivists must find a way to consistently claim two things: first, that Rayan has a general right to privacy; and second, that Sam acts permissibly in reading Rayan's diary in Twin Impersonation, given that he has absolutely no reason not to read the diary.[55] There are two principal ways of combining these claims. On the one hand, perspectivists can specify Rayan's right to privacy in such a way that it no longer covers Twin Impersonation, allowing us to retain the close conceptual connections between rights, obligations, and reasons that are commonly assumed. This leads us to $\text{Privacy}_{\text{Persp}}$, with all of the problems that I outlined above. On the other hand, perspectivists can reject Obligation-Reason and hold that even though Sam is obligated not to read the diary, Sam has (or at least can have) no reason not to read the diary.

Interestingly, what we find here is a dialectic that exactly mirrors the structure of the debate between specificationism and the *pro tanto* view of rights. The only difference is that while the former two are traditionally concerned with the question of whether the conclusive force of rights can be defeated by countervailing considerations, the current dialectic concerns the question of whether the conclusive *and contributory* normative force of rights can be defeated by limitations in the epistemic position of agents. My suggestion is to take seriously these parallels, jettison Obligation-Reason, and extend the category of rights infringements that Thomson makes such a plausible case for. This would allow us to say that universal rights and obligations do indeed cover cases like Twin Impersonation. However, in such cases, acting contrary to rights and their corresponding obligations represents not a violation of them, only a mere infringement.

55 At least not on account of Rayan's privacy. For the sake of simplicity, I shall assume (unrealistically) that there are no other reasons that might speak against reading the diary in this situation.

5.2. Prima Facie Rights and Obligations

On such a view, we must give up the idea that rights and obligations always make a normative contribution in the sense of giving others reasons to comply with their demands.[56] Instead, the view conceives of them as what we may call only *prima facie normative*. In what follows, I briefly lay out what this would entail and then show why it does not force us to give up on the most important features of rights. First of all, let me offer some brief clarifications on my use of the term *prima facie*. I am of course here following W. D. Ross, whose theory of *prima facie duties* is of foundational importance for modern versions of normative pluralism. On the Rossian usage, something being *prima facie x* does not mean that it is only apparently *x*, suggestive of *x*, or something of this sort. Instead, Ross glosses his usage of *prima facie x* as meaning something like "having a tendency to be *x*."[57] Somewhat more precisely, we can say that if something is *prima facie x*, it will be *x* unless certain special circumstances obtain.

Matters are complicated somewhat by the fact that Rossian *prima facie* duties are nowadays commonly assimilated with moral reasons. What it is for an action to be a *prima facie* duty in Ross's sense, many hold, just is for it to be supported by a moral reason.[58] However, this interpretation is meant to make good on the fact that Rossian duties are of a *prima facie requiring* nature, given that duties are the kind of things that generally require. This means that *prima facie* duties only have a tendency to require the relevant option but do not necessarily do so, just as moral reasons plausibly do.

What I suggest is that obligations corresponding to moral rights are in fact *doubly prima facie*. Like Rossian *prima facie* duties, they have a *prima facie* requiring nature, but unlike reasons, they also have a *prima facie contributory* nature. This means that they not only merely have a tendency to require (without necessarily doing so) but that they also merely have a tendency to normatively *count in favor* of the relevant option (without necessarily doing so). Building on the distinction just sketched, we can therefore separate three different views on rights: all-things-considered rights (specificationism), *pro tanto* rights, and *prima facie* rights. These three views are distinguished by progressively more

56 Though even when failing to do so, they will regularly still give rise to other reasons, e.g., reasons for the rights holder to demand justification or explanation. I return to this point in the next section.

57 See Ross, *The Right and the Good*, 28.

58 See Stratton-Lake, "Introduction," xxxiii–xxxviii; Shaver, "Ross on Self and Others"; and Cowan, "Rossian Conceptual Intuitionism," 825.

conditional takes on the normative force of rights. The distinction is illustrated schematically in table 1.

Table 1. Three Different Accounts of Rights

All-things-considered rights (specified)	*Pro tanto* rights	*Prima facie* rights
Actual requirements	*Prima facie* requirements Actual contribution	*Prima facie* requirements *Prima facie* contribution

To state the general idea behind the *prima facie* view more clearly, it is helpful to capture it somewhat more formally. We set off from a need for a nonspecificationist answer to cases like Blood Type, which pushes us toward a rejection of Obligation-Ought and therefore to acceptance of the following principle.

> *Requirement*$_{Prima\ facie}$: If *S* has an obligation *x* to *R* not to ϕ, then *S* ought not to ϕ unless normative considerations more important than *x* favor ϕ-ing.

As we have seen, the most suggestive explanation of why Requirement$_{Prima\ facie}$ is true is one that connects obligations to contributory considerations that jointly determine the overall ought—i.e., normative reasons as commonly understood. To capture our intuitions about Twin Impersonation without any implausible specification, the *prima facie* view of rights and obligations goes beyond this by adding an element of conditionality on the contributory level:

> *Contribution*$_{Prima\ facie}$: If *S* has an obligation *x* to *R* not to ϕ, then *S* has a reason not to ϕ, unless the grounds of *x* are not epistemically accessible to *S*.

As a corollary of Contribution$_{Prima\ facie}$ and a plausible connection between reasons and ought, we are also led to an adjustment of our account of the potentially requiring force of obligations. We must now also accept:

> *Requirement*$_{Prima\ facie}$′: If *S* has an obligation *x* to *R* not to ϕ, then *S* ought not to ϕ unless (1) (epistemically accessible) normative considerations more important than *x* favor ϕ-ing, or (2) the grounds of *x* are not epistemically accessible to *S*.

Contribution$_{Prima\ facie}$ and Requirement$_{Prima\ facie}$′ jointly characterize the *prima facie* view of rights and obligations, allowing for a rejection of Obligation-Reason and a solution to the rights trilemma. By extending the class of permissible rights infringements to cases like Twin Impersonation, we can retain the core perspectivist commitments without slipping into any form of perspectivalism.

5.3. Nonepistemic Restrictions to the Contributory Force of Obligations

It is important to note that the structural features of the *prima facie* view of rights and obligations just laid out do not commit us to a restriction that is exclusively catered to solving the rights trilemma of perspectivism—i.e., to the addition of only an epistemic exception clause to our accounts of both the requiring and contributory force of obligations. Once we go *prima facie*, there is no structural reason to restrict ourselves to only one kind of additional exception clause. While the *prima facie* view of rights does not require that there are further exception classes, it is nonetheless instructive to at least consider potential candidates. To offer but one example, we might also consider adding a clause of inability.

> *Requirement*$_{\textit{Prima facie}}$″: If *S* has an obligation *x* to *R* not to ϕ, then *S* ought not to ϕ unless (1) normative considerations more important than *x* favor ϕ-ing, or (2) the grounds of *x* are not epistemically accessible to *S*, or (3) *S* is unable not to ϕ, or (4) ….

An addition of this inability clause could be motivated by the idea that we can have rights even against those who find themselves unable to properly respect them. For example, a person who enters your house under an irresistible post-hypnotic suggestion might arguably be said to infringe on your rights to property and privacy (although we would likely not want to speak of a *violation* of your rights here). Those who want to uphold the idea that the right to property covers cases of hypnotically forced intrusion would have to add an exception clause like 3 above. After all, the widely accepted principle of Ought Implies Can yields that it is not the case that the hypnotized person ought not to enter your house (since they could not avoid doing so). What is more, if we uphold not only the uncontroversial Ought Implies Can but also the slightly more contested principle of Reason Implies Can, we also arrive at a second class of situations in which rights and obligations lack contributory force.[59] This would then yield a second dimension along which rights and obligations can be said to be doubly *prima facie*.

> *Contribution*$_{\textit{Prima facie}}$′: If *S* has an obligation *x* to *R* not to ϕ, then *S* has a reason not to ϕ, unless (1) the grounds of *x* are not epistemically accessible to *R*, or (2) *S* is unable not to ϕ.

In fact, it is highly likely that Thomson, perhaps the most important defender of an infringement view of rights, would endorse something like the added inability clauses. This becomes clearest in Thomson's writing on self-defense. In

59 For arguments for this principle, see, e.g., Streumer, "Reasons and Impossibility."

her seminal paper on the topic, Thomson argues that it is permissible to defend oneself against what she calls an Innocent Threat—a person who poses a lethal risk without exhibiting any aggressive agency at all.[60] The most commonly cited examples of Innocent Threats involve human projectiles who stand to lethally crush Victim in a way that would not lead to Innocent Threat's own death. Thomson's rights forfeiture account provides an explanation of the permissibility of lethal defensive measures on Victim's part that centrally relies on the ideas that Innocent Threat stands to violate Victim's right to life and that this can be the case even when Innocent Threat does not exhibit any agency at all.[61] As such, and given correlativity (which she endorses), Thomson must allow that Innocent Threat has an obligation not to kill Victim, even though Innocent Threat clearly lacks the ability not to do so.

I do not here want to take a stand on the plausibility of including a global inability clause. Perhaps rights can be infringed only through some kinds of unavoidable behavior and must be specified to exclude others. And perhaps only some rights (e.g., fundamental ones such as the right to life) can be infringed by those who cannot help but do so. Perhaps we should even reject the possibility of rights infringements by those who cannot do otherwise altogether. What I hope has become clear, however, is that there are at least no structural reasons why considerations of epistemic access should be alone in motivating a move toward a *prima facie* view of rights and obligations. Whichever way we ultimately position ourselves, I hope that this highlighting of potential *companions in guilt* to the epistemic exception clause that I propose goes at least some way toward rendering this suggestion more plausible.

6. CONSEQUENCES OF GOING *PRIMA FACIE*

I have sketched a *prima facie* view of rights and obligations that allows perspectivists to uphold both Correlativity and rights universalism. However, the way in which it does so might give rise to a worry: Does the *prima facie* view not seriously denigrate the importance of rights and obligations, especially in situations like Twin Impersonation that cause perspectivism trouble to begin with? Why should rights (and obligations) even interest us if they are not normative *sans phrase*—i.e., normative in the sense of always at least playing a contributory role in determining all-things-considered ought judgments? This

60 See Thomson, "Self-Defense."

61 Thomson, "Self-Defense," 301–3. To yield a *prima facie* view of rights involving an inability clause, we need not share what is perhaps the most controversial element of this claim—namely, that Innocent Threats are guilty of impermissible rights *violations* rather than simply cases of rights infringement.

worry becomes especially salient given the common picture of rights as a kind of normative shield, which I alluded to above. How are rights supposed to properly protect us, one might ask, if they do not necessarily give rise to any reasons for others not to subject us to these outcomes?

I believe that perspectivists who opt for the *prima facie* view can and should face this worry head on. Rights and obligations retain two of their most important features even in situations where they do not make normative contributions in the sense just mentioned. What is more, both of these features are crucial to the sense of protection intuitively provided by rights. These two features are, first, entitlements to demand of others that they respect the rights we have and, second, follow-up duties in the cases of nonrespected rights. Even on the *prima facie* view, rights thus still matter in the way that we intuitively take them to.

6.1. Entitlements to Demand Compliance

Let us dwell in some more detail on the idea of normative protection provided by rights. What rights are for, on this view, is to protect us from being treated in certain ways that are not compatible with our fundamental interests and/or dignity. What is more, rights protect us in such a way that we can actively avail ourselves of this very protection—they give us an entitlement to demand compliance from those against whom we hold the right. My right to privacy involves not only a claim against others that they do not sneak a peek at my most private thoughts but also (via the claim) a power to legitimately demand of them that they do not do so. One might now worry that the *prima facie* view, on which I can also have rights against those who have absolutely no reason not to read my diary, cannot account for this power and thus for this crucial function of rights.

However, this conclusion would be hasty. The *prima facie* view is in fact compatible with a far-reaching entitlement to demand compliance, even in situations in which epistemic limitations obtain that would lead to only a rights infringement, not a violation. The *prima facie* view can account for this because in the problematic situations, the act of demand itself changes the epistemic situation for the addressee. In the very act of citing a right against a previously innocently ignorant party, a right holder can give their addressee crucial new evidence. The demand itself gives the addressee grounds for assuming that there are normatively relevant features of the situation that they have hitherto missed, creating a requirement for them to at least show caution and inquire further before acting against the demand.

To illustrate, imagine a version of Twin Impersonation in which Rayan monitors her room via security camera and, when spotting Sam opening the drawer of her bedside table, simply declares via the intercom, "I have a right that you do not read my diary. You must stop what you are doing right now!" This

might leave Sam confused. After all, for all he knows, he was just given explicit consent by Rayan (impersonated by Riya). Nonetheless, it seems clear that after the announcement from the intercom, Sam can no longer simply permissibly continue with his original plan. After the intervention, his evidential position no longer unequivocally supports a belief in the permissibility of perusing the diary, at least until he has sufficiently clarified the situation. Rayan's demand is therefore legitimate—she does not ask too much of Sam in calling him out, even though prior to her intervention, he had no reason not to proceed.

Some might be given pause by the fact that the very act of intervention creates the situation that supposedly justifies it. Is this not a kind of problematic bootstrapping of justification? I think it is not. For one thing, notice that the legitimacy of the intervention crucially depends on the fact that there actually *is* further evidence available to Sam through a more thorough investigation (calling Rayan on the phone, for example), and that this further evidence ultimately clearly supports a prohibition against Sam reading the diary. To illustrate this necessity, imagine a slightly different scenario. In this alternative case, Rayan makes a similar intervention over the intercom when Sam picks up the latest edition of *Vogue* magazine, which is lying on the coffee table. The magazine, let us assume, contains a saucy story about Rayan's recent appearance on a reality TV show. A nonspecific demand by Rayan against Sam to stop reading the magazine on account of her privacy might similarly startle Sam and temporarily stop him from reading. In this case, however, investigating further and reasoning correctly would not lead him to the conclusion that he ought not to read the magazine. After all, our right to privacy does not plausibly give us a claim against others that they do not read unfavorable news articles about our voluntary appearances on public television. For these reasons, Rayan's demand would be illegitimate in this alternative scenario and not create a robust requirement for Sam to permanently desist from reading.

What is more, there are other situations that take a structurally similar form to the creation of reasons through demands on the *prima facie* view. For example, a legitimate authority may issue a command by saying something like "You are required to go home immediately," although it is precisely the act of demand by the authority itself that makes the relevant action required. And perspectivists must account for similar apparently "self-fulfilling prophecies" in the more general phenomenon of advice, as a better-informed party can truly offer guidance of the sort "you really ought to ϕ" even in situations in which the reasons for ϕ-ing are not yet available to the advisee.[62]

62 However, perspectivists may have to jump through some hoops to do so. For possible solutions, see Kiesewetter, "'Ought' and the Perspective of the Agent"; and Lord, "Acting for the Right Reasons, Abilities, and Obligation."

Contrary to first appearances, the *prima facie* view of rights is therefore, after all, capable of accounting for one of the most important conceptual features of rights—generally, rights holders are entitled to demand compliance from both would-be violators and would-be infringers equally.

6.2. Follow-Up Duties

The second feature of rights pertains to the abovementioned follow-up duties they generate. An important way in which rights matter is that the normative situation changes significantly when they are not respected. Rights violators usually have weighty duties of apology and repair to those they have wronged. However, as I argued above, even rights whose corresponding obligations are defeated by countervailing considerations are not thereby rendered wholly normatively inert in the wake of contrary action. In such cases of rights infringement, there is likewise a *remainder*, giving rise to duties of explanation, compensation, and, in certain cases, also apology.

The same plausibly applies to obligations defeated by epistemic inaccessibility. If I permissibly break into your cabin, whether it is to avoid freezing to death or because I innocently believe it is my newly acquired property after some pranksters carefully exchanged the signage, I incur duties of explanation and, plausibly, compensation for damages caused. In both cases, the fact that you have a right to the cabin offers a straightforward and convincing explanation of these follow-ups. Had you obtained the cabin by illicit means or recently sold the cabin to a third party, I would have no such duties toward you. Again, we can see that rights matter even without directly giving rise to reasons for action.

Now it might be objected that the follow-up duties that the perspectivist can allow for are themselves much less robust than one might hope. After all, their obtaining in turn must depend on the evidential situation of the infringing party. Sure, if I later find out that the mislabeled cabin was yours after all, I will be required to explain myself and potentially indemnify you (at least to some degree). But until I obtain that information, no follow-ups are incurred in the wake of my unwitting but nonetheless real infringement of your right. What use is the ability to draw on rights in explaining remainders then, if all we get is such an instable result?

It is true that on a perspectivist picture, follow-up duties must be epistemically constrained in just this way. However, that does not mean that the *prima facie* view's ability to draw on rights infringement even in cases such as Twin Impersonation or Feinberg's cabin is worthless. In the good case in which the agents do find out about their mistake, the rights infringement plays a crucial role in the most elegant and straightforward explanation of the normative remainder. A form of rights perspectivalism, which cannot draw on any sense

of rights infringement here, is left in a poorer explanatory position. What is more, even in the bad case, the *prima facie* view can again help us explain why it is appropriate for the victim to demand compliance with follow-up duties. For reasons laid out in the previous subsection, victims have an entitlement to demand explanation, indemnification, and so on, even before the infringing party finds out about their previous evidential shortcomings. Since in doing so they will change the evidential situation of their addressees, rights holders whose rights were infringed can stand on these very rights in the aftermath to demand that things be put right. In this way, rights retain additional value to them even if their reason-giving force is normatively constrained in the way perspectivists must claim.

6.3. Softening the Blow by Disambiguation?

Although the two features just sketched allow for *prima facie* rights to retain a normatively important role, it cannot be denied that the move of going *prima facie* and denying Obligation-Reason nonetheless represents a substantial step away from most people's intuitive conceptual starting points regarding rights. In response to this, one might be tempted to further soften the blow by adopting a disambiguationist response.[63] For one thing, one could follow some perspectivists in making a distinction between reasons for ϕ-ing that there are for an agent *S*, on the one hand, and reasons that an agent *S* has for ϕ-ing, on the other.[64] This would allow perspectivists to uphold the intuitive verdict that there are reasons for agents like Sam to comply with their obligations—reasons that can, for example, also be drawn upon by third parties offering advice. At the same time, it allows us to also retain the crucial perspectivist claim that the absence of actual consent by Rayan does not play a role in what Sam ought to do, all things considered. For even though the absence of consent explains why *there are* weighty reasons not to read the diary, a perspectivist would claim that these are not reasons *that Sam has*. On the view under consideration, only reasons possessed by *S* can impinge on the question of what *S* ought to do, and *S*'s perspective is a core limiting factor on what reasons *S* can possess. While we still must deny what we may call Obligation-Reason$_{\text{Possessed}}$, we can therefore at least uphold Obligation-Reason$_{\text{Existing}}$. This is of course but one way of pursuing the more general strategy of capturing opposing intuitions (or at least coming closer to doing so) by disambiguating some of our talk of reasons or ought. Those not taken with the quasi-possessive notion of having reasons

63 I thank Benjamin Kiesewetter for pressing me to address this issue in more detail.

64 See Lord, *The Importance of Being Rational*.

just sketched might, for example, simply distinguish between subjective and objective reasons, assigning different incompatible roles to each of them.[65]

Whichever disambiguationist strategy one might opt for, it is important not to overstate their overall potential. Whichever way we choose to draw the lines between more objective and more subjective versions of normative concepts, the foregoing discussions of the rights trilemma and the failures of rights perspectivalism clearly bear out one thing: the most familiar notion of rights—the one we are used to employing in everyday discourse—operates squarely on the objective side of things. The most important notion of rights is thus at a remove from the most important notions of ought, reason, etc., which any perspectivist worth their salt must firmly locate on the nonobjectivist side of the divide. There might be some notion of subjective rights to keep the perspectivist reasons and oughts company on their side of the chasm, and there may also be an objective (or nonpossessed) kind of reason joining objective rights on the other, but the intuitively more important notions of rights and reasons remain at a distance. Since attempts to forcibly draw the notion of rights fully over onto the nonobjective side fail (i.e., rights perspectivalism fails), even a committed perspectivist disambiguationist must admit that Obligation-Reason does not hold when keeping fixed the intuitively more important versions of each notion.

Given this result, it is all the more important to keep in focus that there are bridges, one might say, between strictly objective notions of rights and nonobjective notions of reasons and oughts. The two most important of these are the abovementioned entitlements to demand compliance and the possibility of follow-up duties. Nonetheless, it would be remiss to deny that the perspectivism-friendly notion of rights that I have sketched represents a substantial rethinking of moral claim rights—one that requires us to give up some widely held beliefs about their normative force. Disambiguationism provides no way out of this conclusion. However, since others have recently argued that a thorough rethinking of moral claim rights is long overdue for other, unrelated reasons, this may finally not even end up as an unwelcome result, as long as we can retain enough of the most important features that give us reason to engage in rights discourse in the domain of morality in the first place.[66]

65 See Schroeder, "Having Reasons."

66 For a recent argument in favor of a wholesale reconsideration of moral claim rights, see Valentini, "Rethinking Moral Claim Rights."

7. CONCLUSION

I have argued that the rights trilemma for perspectivism, combined with the untenability of rights perspectivalism, forces any committed perspectivist to thoroughly reconceptualize the normative relevance of rights and obligations. Instead of having rights universally give rise to (moral) reasons that can affect what those against whom we hold rights ought to do, perspectivists can only allow for the *prima facie* normativity of rights and obligations and must therefore concede that rights and obligations sometimes fail to give obligation holders any reason for compliance at all.

Though this represents a relatively radical reinterpretation of the normative relevance of rights, I believe it need not be one that relegates rights and their corresponding obligations to a position of unimportance. I have outlined two important roles played by rights and obligations to then show how they can play both of these roles even when they do not correspond to reasons for complying with them due to epistemic defeat. Even when they are only *prima facie* normative, rights still offer their holders a form of protection that they are right to value. This, of course, is far from a conclusive case for the *prima view* of rights. Many philosophically interesting questions remain open. If what I have argued is correct, however, the relative merits of the *prima facie* view of rights when compared to the other horns of the rights trilemma give perspectivists good reason to carefully inquire into these questions and develop the *prima facie* view in greater detail.[67]

Humboldt University of Berlin
daniele.bruno@hu-berlin.de

REFERENCES

Andrić, Vuko. *From Value to Rightness: Consequentialism, Action-Guidance, and the Perspective-Dependence of Moral Duties*. Routledge, 2021.

67 Material from this article has been presented at the Humboldt Normativity Conference and at colloquia in Göttingen, Erlangen, Munich, Berlin, and Bielefeld. I have profited immensely from the helpful exchanges on all of these occasions and am grateful to both organizers and participants for making them possible. Thanks to Francesca Bunkenborg, Benjamin Kiesewetter, Erasmus Mayr, Thomas Schmidt, Holmer Steinfath, Laura Valentini, and two anonymous referees for *JESP* for their helpful feedback on earlier versions of this article. Work on this article was supported by Deutsche Forschungsgemeinschaft (project number 503995582).

———. "Objective Consequentialism and the Rationales of 'Ought' Implies 'Can'." *Ratio* 30, no. 1 (2017): 72–87.

Botterell, Andrew. "In Defence of Infringement." *Law and Philosophy* 27, no. 3 (2008): 269–92.

Camp, Elizabeth. "Perspectives and Frames in Pursuit of Ultimate Understanding." In *Varieties of Understanding*, edited by Stephen R. Grimm. Oxford University Press, 2019.

Cowan, Robert. "Rossian Conceptual Intuitionism." *Ethics* 127, no. 4 (2017): 821–51.

Dancy, Jonathan. *Practical Reality*. Oxford University Press, 2000.

Dworkin, Ronald. *Taking Rights Seriously*. Harvard University Press, 1977.

Feinberg, Joel. *Rights, Justice, and the Bounds of Liberty: Essays in Social Philosophy*. Princeton University Press, 1980.

———. "Supererogation and Rules." *Ethics* 71, no. 4 (1961): 276–88.

Fox, Philip. "Revisiting the Argument from Action Guidance." *Journal of Ethics and Social Philosophy* 15, no. 3 (2019): 222–54.

Frederick, Danny. "*Pro Tanto* Versus Absolute Rights." *Philosophical Forum* 45, no. 4 (2014): 375–94.

Griffin, James. "Discrepancies Between the Best Philosophical Account of Human Rights and the International Law of Human Rights." *Proceedings of the Aristotelian Society* 101, no. 1 (2001): 1–28.

Hart, H. L. A. "Are There Any Natural Rights?" *Philosophical Review* 64, no. 2 (1955): 175–91.

Hinsch, Wilfried, and Markus Stepanians. "Human Rights as Moral Claim Rights." In *Rawls's Law of Peoples: A Realistic Utopia?* edited by Rex Martin. Blackwell, 2009.

Hohfeld, Wesley Newcomb. *Fundamental Legal Conceptions*. Edited by Walter Wheeler Cook. Yale University Press, 1919.

Jackson, Frank. "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection." *Ethics* 101, no. 3 (1991): 461–82.

Kiesewetter, Benjamin. *The Normativity of Rationality*. Oxford University Press, 2017.

———. "'Ought' and the Perspective of the Agent." *Journal of Ethics and Social Philosophy* 5, no. 3 (2011): 1–25.

———. "*Pro Tanto* Rights and the Duty to Save the Greater Number." In *Oxford Studies in Normative Ethics*, vol. 13, edited by Mark Timmons. Oxford University Press, 2023.

———. "What Kind of Perspectivism?" *Journal of Moral Philosophy* 15, no. 4 (2018): 415–43.

Littlejohn, Clayton. "Is Justification Just in the Head?" In *Contemporary Debates*

in Epistemology, 3rd ed., edited by Matthias Steup, Blake Roeber, John Turri, and Ernest Sosa. Wiley-Blackwell, 2024.

Lord, Errol. "Acting for the Right Reasons, Abilities, and Obligation." In *Oxford Studies in Metaethics*, vol. 10, edited by Russ Shafer-Landau. Oxford University Press, 2015.

———. *The Importance of Being Rational*. Oxford University Press, 2018.

Nozick, Robert. *Anarchy, State, and Utopia*. Basic Books, 1974.

Oberdiek, John. "Lost in Moral Space: On the Infringing/Violating Distinction and Its Place in the Theory of Rights." *Law and Philosophy* 23, no. 4 (2004): 325–46.

———. "Specifying Rights Out of Necessity." *Oxford Journal of Legal Studies* 28, no. 1 (2008): 127–46.

Owens, David. *Shaping the Normative Landscape*. Oxford University Press, 2012.

Parfit, Derek. *On What Matters*. Vol. 1. Oxford University Press, 2011.

Raz, Joseph. *The Morality of Freedom*. Clarendon Press, 1986.

———. *Practical Reason and Norms*. Oxford University Press, 1999.

Regan, Tom. *Animal Rights, Human Wrongs: An Introduction to Moral Philosophy*. Rowman and Littlefield, 2003.

Rettig, Cristián. "Rights and Practical Reasoning: A Practical View on the Specificationism vs. Generalism Debate." *Journal of Value Inquiry* (forthcoming). Published online ahead of print, August 12, 2023. https://doi.org/10.1007/s10790-023-09946-5.

Rettig, Cristián, and Giulio Fornaroli. "Conflicts of Rights and Action-Guidingness." *Ratio Juris* 36, no. 2 (2023): 136–52.

Rosen, Gideon. "Culpability and Ignorance." *Proceedings of the Aristotelian Society* 103 (2003): 61–84.

Ross, W. D. *The Right and the Good*. Edited by Philip Stratton-Lake. Clarendon Press, 2002.

Schauer, Frederick. "A Comment on the Structure of Rights." *Georgia Law Review* 27 (1992): 415–34.

Scanlon, T. M. *What We Owe to Each Other*. Harvard University Press, 1998.

Schmidt, Thomas. "The Balancing View of Ought." *Ethics* 134, no. 2 (2024): 246–67.

Schroeder, Mark. "Having Reasons." *Philosophical Studies* 139, no. 1 (2008): 57–71.

Shafer-Landau, Russ. "Specifying Absolute Rights." *Arizona Law Review* 37 (1995): 209–25.

Shaver, Robert. "Ross on Self and Others." *Utilitas* 26, no. 3 (2014): 303–20.

Sliwa, Paulina. "Making Sense of Things: Moral Inquiry as Hermeneutical Inquiry." *Philosophy and Phenomenological Research* 109, no. 1 (2023): 117–37.

Stratton-Lake, Philip. "Introduction." In Ross, *The Right and the Good*.
Streumer, Bart. "Reasons and Impossibility." *Philosophical Studies* 136, no. 3 (2007): 351–84.
Thomson, Judith Jarvis. *The Realm of Rights*. Harvard University Press, 1990.
———. "Self-Defense." *Philosophy and Public Affairs* 20, no. 4 (1991): 283–310.
———. "Self-Defense and Rights." In *Rights, Restitution, and Risk: Essays in Moral Theory*, edited by William Parent. Harvard University Press, 1986.
Valentini, Laura. "Rethinking Moral Claim Rights." *Journal of Political Philosophy* 31, no. 4 (2023): 433–51.
Wallace, R. Jay. *The Moral Nexus*. Princeton University Press, 2019.
Way, Jonathan, and Daniel Whiting. "Perspectivism and the Argument from Guidance." *Ethical Theory and Moral Practice* 20, no. 2 (2017): 361–74.
Wenar, Leif. "Rights." In *Stanford Encyclopedia of Philosophy* (Spring 2023). https://plato.stanford.edu/archives/spr2023/entries/rights/.
Wellman, Christopher Heath. "On Conflicts Between Rights." *Law and Philosophy* 14, nos. 3/4 (1995): 271–95.
Williams, Bernard. *Ethics and the Limits of Philosophy*. Routledge, 2006.
Zimmerman, Michael J. *Ignorance and Moral Obligation*. Oxford University Press, 2014.
———. *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge University Press, 2008.

Journal of Ethics and Social Philosophy
VOL. 30, NO. 3 · MAY 2025
https://doi.org/10.26556/jesp.v30i3.3530

WHY WE SHOULD UNBUNDLE THE POLICE

Lauren Lyons

THE ALARMING RECURRENCE of unjustified killings by police highlights systemic issues that should be deeply concerning to us all. Beyond excessive use of force, the police treat marginalized people in disproportionately harmful ways that reflect and perpetuate endemic injustice; they respond inappropriately to complex social and public health problems like homelessness, addiction, and mental illness, risking harmful escalation and exacerbating underlying issues.[1] Police culture tends towards cynical authoritarianism, adopting an "us-versus-them" mentality that positions (at least a subset of) citizens as adversaries.[2] All of this has resulted in severely diminished public trust in the police, fraught police-community relations, and rising skepticism of the legitimacy of policing institutions.[3]

Public outcry over these problems has catalyzed the ongoing Black Lives Matter movement. The police murder of George Floyd was followed by mass protests in the summer of 2020, and since then, there has been widespread public debate on how to mitigate police violence and the distrust it engenders. Some call for incremental reforms, like changing laws and policies governing police use of force or strengthening misconduct reporting and decertification processes.[4] Others demand that we reimagine the role of policing in our institutional landscape, reallocating powers, resources, and responsibilities from the

1 On police brutality, see Zimring, *When Police Kill*; Ralph, *The Torture Letters*; the *Washington Post* police shootings database (2015–2024), https://www.washingtonpost.com/graphics/investigations/police-shootings-database/; and Mapping Police Violence, "2024 Police Violence Report." On race and policing, see Butler, *Chokehold*; Ritchie, *Invisible No More*; and Davis, *Policing the Black Man*. On policing poverty and the effects of policing on people with addiction, housing insecurity, and mental illness, see Macaré et al., *Who Do You Serve, Who Do You Protect?*; Clifford, *Policing, Mental Illness, and the Media*; Wacquant, *Punishing the Poor*; and Vitale, *City of Disorder*.

2 See Balko, *Rise of the Warrior Cop*; and Sierra-Arévalo, *The Danger Imperative*.

3 See Bell, "Police Reform and the Dismantling of Legal Estrangement"; Goldsmith, "Police Reform and the Problem of Trust"; and Brown and Lloyd, "Black Americans Less Confident, Satisfied with Local Police."

4 For an overview of the state of these sorts of reforms in the United States, see Subramanian and Arzy, "State Policing Reforms Since George Floyd's Murder."

police to other institutions.[5] The goal of this paper is to refine and defend this reallocative demand, which I refer to as the *unbundling proposal*.[6]

There has been a promising uptick in philosophical discussions of policing in recent years. Some focus on principles to guide police conduct, often drawing on theories of self-defense and professional ethics.[7] Philosophers also propose measures to address police misconduct such as expanding legal statutes to outlaw harmful tactics, revoking the licenses of bad actors, providing reparations to victims of police violence, implementing self-evaluation and evidence-based improvements to departmental policy, restructuring police departments, broadening police participation in harm reduction and other forms of nonviolent order maintenance, and avoiding tactics that heighten the risk of illegitimate policing.[8]

These strategies, especially when combined, can improve policing. Rather than a discussion of their comparative merits and disadvantages, I present and defend an alternative ameliorative approach. The unbundling proposal asks not how police should act but rather what the scope of policing should be: Which situations require police presence? In the ethics of war, we distinguish between *jus in bello* (the ethics of conduct in war) and *jus ad bellum* (the ethics of whether war is justified). The unbundling proposal addresses an issue that is analogous to *jus ad bellum* considerations: when police should be deployed (instead of how they should behave).[9] This approach complements rather than

5 Some organizations that support the reallocative demand include MPD150 (Minneapolis), Critical Resistance (international), Project Nia (Chicago, New York), Interrupting Criminalization (United States), and the Oakland Power Project (Oakland). Some influential defenses of police abolition and reallocative measures include Kaba, "Yes, We Mean Literally Abolish the Police"; Vitale, *The End of Policing*; Maher, *A World Without Police*; Kaba and Ritchie, *No More Police.*; and Purnell, *Becoming Abolitionists*.

6 The term 'unbundling' was coined by musician and activist Trevor McFedries (Thompson, "Unbundle the Police") and has entered common usage. Some illuminating discussions of unbundling include Thompson, "Unbundle the Police"; Thacher, "Shrinking the Police Footprint"; and Friedman, "Disaggregating the Police Function."

7 See Hunt, "Policing, Brutality, and the Demands of Justice"; Monaghan, "The Special Moral Obligations of Law Enforcement" and *Just Policing*; and Page, "Defensive Killing by Police."

8 On expanding legal statutes and revoking the licenses of bad actors, see Hunt, "Policing, Brutality, and the Demands of Justice"; and Jones, "Police-Generated Killings." On reparations for police violence, see Page, "Reparations for Police Killings." On reforms to departmental policy, see Monaghan, "Legitimate Policing and Professional Norms." On restructuring police departments, see Monaghan, *Just Policing*, ch. 8. On police participation in harm reduction, see Monaghan, "Broken Windows, Naloxone, and Experiments in Policing." On avoiding tactics that heighten the risk of illegitimate policing, see Monaghan, *Just Policing*, chs. 3–6.

9 Alice Ristroph draws several helpful analogies between the ethics of war and criminal law, arguing that philosophical and legal approaches to punishment should adopt a *jus in bello*–inspired principle aimed at limiting the violence of punishment—what she calls *jus*

conflicts with many proposed reforms, but it also addresses a broader and less examined issue. Moreover, despite substantial public support, there has been no sustained discussion of unbundling in analytic ethics and political philosophy, and the attention the proposal has received is largely critical.[10]

The unbundling proposal is connected closely to movements to defund and eventually abolish the police. The slogan "defund the police" really means "defund and refund," with activists calling for cutting police funding *and* reallocating it to other nonpolice institutions and community organizations.[11] As such, "defund, refund" is one public finance-focused component of the broader unbundling proposal. For abolitionists, unbundling and other measures that reduce the scope and power of the police are critical steps toward ultimately dismantling the institution. Though I am not defending abolition here, the discussion should (1) clarify the practical action strategy of police abolitionists and (2) offer a more robust and appealing picture of the defund demand.[12]

The structure of this paper is as follows. In section 1, I present the unbundling proposal, identifying the specific dimensions of policing that proponents argue should be unbundled and reallocated. There I also discuss the definition of policing upon which unbundling is based. Then, I present a novel set of normative arguments for unbundling that reflect various rationales emanating from policing-critical social movements. The case for unbundling is strongest if we take them in tandem. The first two arguments (section 2) draw on principles

in poena. She does not explicitly link *jus ad bellum* and the scope of policing introduced here, but she does discuss how interpersonal morality-focused approaches to punishment tend to obscure the role of the state as the agent of violence. See Ristroph, "Just Violence."

10 Joseph Heath, Luke Hunt, and Jake Monaghan critically discuss proposals related to unbundling. Heath warns of the "deleterious consequences" of removing police from order maintenance, citing potential increases in crime and police violence ("The Challenge of Policing Minorities in a Liberal Society"). Hunt raises similar concerns with "reallocative policing," though he is open to piecemeal reform ("The Limits of Reallocative and Algorithmic Policing"). He questions how reallocative measures would work given high gun ownership and the *definition* of policing upon which reallocation is based. Critiquing abolitionism, Monaghan poses a similar challenge about defining policing, noting that all alternatives to policing involve some form of social control and thus policing, meaning "the risk of unjust policing is always with us" (*Just Policing*, 17). Elsewhere, Monaghan advocates restructuring police departments to separate law enforcement from order maintenance (*Just Policing*, ch. 8). My proposal extends this idea, arguing that police should not be involved at all in routine order maintenance.

11 Kaba and Ritchie, *No More Police*; and Vitale, *The End of Policing*.

12 Daniel Fryer argues that police abolitionism does not "provide a structured and clear vision of radical change" ("Idealizing Abolition," 558). My sense is that while some proponents talk about abolitionism in terms of grand idealizations, others focus on concrete ways that we can build "new non-police institutions" (561). One goal of this discussion is to shed light on these tangible alternatives.

of institutional design. I argue first that we should unbundle policing because public institutions with violent capacities should have narrow mandates; nonviolent, noncoercive responses to social problems should be the default. I then claim that unbundling constitutes a better distribution of epistemic labor. Catchall order-maintenance policing is epistemically overdemanding, while more narrowly defined roles foster better expertise and outcomes. The argument in section 3 centers on the effects of policing in unequal societies with historical injustice—specifically, how policing disproportionately burdens Black people, other people of color, and members of marginalized groups, driving structural injustice. I aim to reconstruct one argumentative thread that leads us from (1) these unfair effects to (2) the unbundling proposal. In doing so, I address the broader question of what forms of solutions are appropriate when institutions are infected with injustice, suggesting that in this case and others, justice-undermining effects require us to turn towards extra-institutional, reallocative measures. My hope is that the paper will be interesting for skeptics *and* advocates of unbundling and related proposals, adding some clarity to divisive debates and expanding the library of solutions to the pressing problems with policing defended within philosophy.

1. THE UNBUNDLING PROPOSAL

Though cutting police powers, responsibilities, and funding has gained prominence in recent social movements, it has a long history in Black radical and abolitionist organizing.[13] Today, unbundling and police abolitionist movements are vast and varied.[14] I present and defend one version of the unbundling proposal, reflecting key demands from these movements, in order to set a clear target for arguments and objections. To understand unbundling, we must first consider current policing practices, as they set a baseline for reallocative measures. In his seminal book *The Ethics of Policing*, John Kleinig argues that (contrary to popular opinion) police work is best characterized as "social

13 Historian Robin Kelley describes how "abolishing the police is not the brainchild of some extreme left-wing think tank," noting that the Black Panther Party was formed "to monitor police violence, to create community-based models of public safety, and to provide for the social needs of Black communities where the state failed" ("What Abolition Looks Like, From the Panthers to the People").

14 There are a variety of organizations around the world, especially in the United States, working on unbundling, defunding, and abolishing police. Some include MPD150 (Minneapolis), Critical Resistance (international), Project Nia (Chicago, New York), Interrupting Criminalization (United States), and the Oakland Power Project (Oakland).

peacekeeping," as opposed to crime fighting.[15] He finds that most of police officers' time is spent doing non-crime-related social service activities such as "intervention in family crises, searching for lost children, rescuing animals, directing traffic, supervising crowds, visiting schools, assisting the elderly, and so on—or in various administrative tasks."[16] In his view, while police are the "repositories of coercive power," their central role is "to ensure or restore peaceful order." Decades of empirical research support this, finding that police indeed spend most of their time doing routine, catchall order maintenance.[17] A 2020 *New York Times* analysis by Jeff Asher and Ben Horwitz of police activity in three jurisdictions illustrates this characterization.[18] Using police dispatch data, they found that noncriminal calls dominated the workload of officers: they made up about 37 percent of calls in New Orleans, Louisiana; 37 percent in Montgomery County, Maryland; and 32 percent in Sacramento, California. By contrast, violent crime calls accounted for around 4 percent in each city, while the remaining share was split among traffic incidents, property crimes, proactive patrols, medical or other assistance, and miscellaneous "other crimes."

The unbundling proposal questions the social peacekeeper model of policing that is the status quo. Requiring police to be the default response to diverse and complex social problems drives police violence and incompetence. Because of this, we should carve off many of these responsibilities. Proposals about what sorts of powers and responsibilities should be reallocated vary, though proponents tend to focus on reallocate measures in five areas, all of which offer paths towards unbundling.[19] My aim is not to offer a comprehensive positive proposal about how we can maintain social order without police but rather to point to some responsibilities that can be plausibly reallocated from police to other institutions. Determining the details about how those other institutions should operate requires attentiveness to context-specific concerns as well as experimentation and revision.

15 Kleinig, *The Ethics of Policing*.

16 Kleinig, *The Ethics of Policing*, 23. Brandon del Pozo also discusses this motley bundle of police duties (*The Police and the State*, 10–12).

17 See Ratcliffe, "Policing and Public Health Calls for Service in Philadelphia"; Webster, "Police Task and Time Study"; Wuschke et al., "What Do Police Do and How Do They Do It?"; and Bittner, "Florence Nightingale in Pursuit of Willie Sutton" and "The Police on Skid Row."

18 Asher and Horwitz, "How Do the Police Actually Spend Their Time?"

19 Some sources upon which this version of unbundling is based include Friedman, "Disaggregating the Police Function"; Kaba, "Yes, We Mean Literally Abolish the Police"; Kaba and Ritchie, *No More Police*; Karma, "4 Ideas to Replace Traditional Police Officers"; Thompson, "Unbundle the Police"; and Vitale, *The End of Policing*.

The first area of intervention is to institute social service–based crisis response. Police are currently responsible for addressing the downstream failures of other social systems, in particular, managing the complex and interrelated problems of serious mental illness, addiction, and homelessness.[20] These issues should not be within the purview of the police but rather managed by community organizations and public institutions more narrowly trained and equipped to deal with them. Social service–based crisis response programs have been instituted throughout the United States and elsewhere. For instance, in Eugene, Oregon, many 911 calls are directed to a program called Crisis Assistance Helping Out on the Streets (CAHOOTS). This publicly funded program has handled calls related to homelessness, addiction, disorientation, and serious mental illness since 1989.[21]

A second intervention tasks trained civilian de-escalators to (1) intervene in minor disputes and (2) promote community safety through patrols. Police presently deal with a variety of minor conflicts related to noise, pets, trespassing, nonviolent arguments, and so on. Domestic violence also comprises a substantial proportion of emergency calls.[22] Police are also present in many schools, where their role has expanded beyond responding to serious threats to include addressing routine behavioral issues. Evidence shows that police presence in schools tends to increase suspensions, expulsions, and student arrests—part of the broader phenomenon of the criminalization of school discipline that disproportionately affects marginalized students.[23]

In a variety of cases, mediation and de-escalation by unarmed trained professionals without the coercive tools of the criminal law are preferable to potentially escalatory police responses. In some models, community safety

20 Thacher, "Shrinking the Police Footprint."

21 CAHOOTS dispatches two-person medical teams equipped to deliver "crisis intervention, counseling, mediation, information and referral, transportation to social services, first aid, and basic-level emergency medical care." Of an estimated 24,000 CAHOOTS calls in 2019, only 311 required police backup. See the case study on CAHOOTS in Vera Institute of Justice, "Behavioral Health Crisis Alternatives." CAHOOTS is funded by the police budget, however, which may be objectionable for some proponents of unbundling. Many places have adopted similar models. San Francisco has opted to dispatch unarmed civilian responders in noncriminal matters such as neighbor disputes, calls about homeless people, and school discipline issues. Albuquerque has created a new category of first responder (beyond police, fire, and EMTs) to dispatch in noncriminal emergencies. Both models are described in Friedman, "Disaggregating the Police Function."

22 Friedman, "Disaggregating the Police Function," 952.

23 See Weisburst, "Patrolling Public Schools"; Mowen and Brent, "School Discipline as a Turning Point"; and Sorensen et al., "Making Schools Safer and/or Escalating Disciplinary Response."

professionals can issue citations but are unarmed and do not have the power to arrest.[24] Police also dedicate significant time to patrolling areas where there is a high risk of violence and to monitoring events like concerts and protests; this work could be reallocated to civilian de-escalators as well.

A third intervention is to institute a civilian traffic patrol system. Police in the United States make over twenty million traffic stops per year.[25] Some of these interactions escalate and have devastating consequences. According to a report by Mapping Police Violence, in 2024, 154 people were killed by police following traffic violations.[26] One way to reduce the incidence of violence in routine traffic interactions is to decouple traffic enforcement and the criminal legal system. Many jurisdictions have systems like this. One example is Highways England, which deploys unarmed traffic officers in nonpolice vehicles to enforce traffic laws through civil means. As described by Barry Friedman, the "law can be brought to bear without force anywhere nearby."[27]

Finally, any reasonable version of the unbundling proposal involves a class of specialized violence responders who are trained and equipped to deal with serious and violent emergencies. In some situations, it is necessary for them to be armed. From an abolitionist perspective, these institutions ought to be distinct from the police, created from the ground up with different training, rules, and procedures. On a more moderate unbundling view, we should do away with most of the present police roles, personnel, and responsibilities, shrinking the size of police departments so that they only intervene in instances where force is potentially required—in other words, reallocating many police powers and responsibilities while keeping some of our present institutional architecture in place.[28]

24 Civilian de-escalation and patrol programs have been successfully implemented in France, the Netherlands, the United Kingdom, Australia, and South Africa. See Gray, "Community Safety Workers."

25 For this data and more, see the Stanford Open Policing Project, https://openpolicing.stanford.edu/.

26 Mapping Police Violence, "2024 Police Violence Report."

27 Friedman, "Disaggregating the Police Function," 960.

28 Currently, 911 dispatchers decide whether to send police, firefighters, medical services, or other responders, often erring on the side of caution by deploying multiple services. Expanding the list of dispatch options by unbundling may complicate these initial decisions. To address this, some cities have introduced nonemergency lines to ease the load on 911 dispatch, and additional dispatcher training can help them recognize situations suited to alternative responders. For example, in Eugene, dispatchers receive training to identify nonviolent cases with mental health components, routing them to CAHOOTS. As with current dispatch systems, mistakes are inevitable. But while sorting calls is challenging,

Proponents of unbundling emphasize that these reallocative measures must be installed alongside increased investment in upstream social support and violence prevention. This dimension of unbundling, unlike the others, does not question police responsibilities; rather, it asks that we take the complex social causes of violence and antisocial behavior seriously and invest in nonpunitive ways to prevent them before they occur. Most concretely, upstream crime prevention entails increased resources for health care, housing, mental health and addiction support, neighborhood improvement, community organizations, and education to reduce social strain and harmful behavior. The scholarly consensus among criminologists and sociologists is that these upstream social reforms effectively curb crime and other forms of violence.[29] Broadly implementing them may render much of what is presently police work "obsolete" in the long term since it will reduce the incidence of problems that demand responses in the first place, as framed in Angela Davis's work on prison abolition.[30]

You should now have a more concrete sense of what the unbundling proposal involves. Using the catchall order-maintenance model of policing as a

---

investing in specialized dispatcher training for new response options is more efficient than broadly training police to handle every type of crisis.

29 There is broad scholarly consensus that crime rates are negatively correlated with levels of welfare assistance and participation. See, e.g., Deshpande and Mueller-Smith, "Does Welfare Prevent Crime?"; Foley, "Welfare Payments and Crime"; and Hannon and DeFronzo, "Welfare and Property Crime." Improvements to the built environment (e.g., greening vacant lots, improving lighting) are also effective. See Eck and Guerette, "Place-Based Crime Prevention." There is also evidence that supports the preventative capacities of education and youth programs (especially youth employment programs). See Modestino, "How Do Summer Youth Employment Programs Improve Criminal Justice Outcomes, and for Whom?" Moreover, health care and mental health/addiction treatment access is correlated with lower crime rates. See Bondurant et al., "Substance Abuse Treatment Centers and Local Crime"; and Vogler, "Access to Health Care and Criminal Behavior." Measures to increase access to affordable housing and reduce neighborhood segregation are also effective. See Chyn, "Moved to Opportunity"; and Freedman and Owens, "Low-Income Housing Development and Crime." Finally, investment in community organizations is correlated with significant reductions in violent crime. See Sharkey, "Why Do We Need the Police?"

30 Davis, "Are Prisons Obsolete?" Another related approach is to restructure existing institutions so as to reduce the necessity of police response, as suggested by David Thacher ("Shrinking the Police Footprint"). Thacher discusses the case of Paducah, Kentucky, where at one point, nearly one in every seven calls handled by the police were initiated by two Walmart stores (largely for shoplifting). In response to this, the police met with store managers and encouraged them to reduce opportunities for shoplifting. Thacher argues that sometimes we should "force recalcitrant institutions to take more responsibility for their own problems" (75). The background view is that police step in when other institutions fail.

baseline, implementing these five interventions would dramatically reduce the size and scope of policing, thus shrinking its institutional footprint.

1.1. Unbundling and the Definition of Policing

Before turning to arguments for unbundling, I address an important conceptual issue. The idea that we should reallocate powers and responsibilities from police to other institutions relies implicitly on a definition of policing. What then, precisely, is the most suitable definition?[31] In common language, we use the word 'police' in various ways. It is intelligible to say, "Instead of police, social workers should respond to mental health crises." And for someone to respond, "Well, social workers are still police!" There are two senses of policing evoked in this exchange.

One is a narrow, formal, and institutional sense of Policing (which will be denoted with an uppercase *P* for clarity in this section), whereby Policing is defined with respect to a formal, specified role within our legal and political system. Laws and policies determine the mandate and official capacities of the Police. Some central distinctive powers include interrogation, search, issuing summons, and arrest. Institutions confer the title of Police to government officials who undergo relevant training and have these capacities. Framed otherwise, Police are just those who have the Policing role within law enforcement agencies, which include local police departments, sheriff's offices, state police, and highway patrol. On this definition, we may claim, "Well, they aren't really Police; they are just mall security."

Alternatively, we may adopt a broad, informal, and practice-based understanding of policing, evoked in the claim that "social workers still police." This definition of policing (denoted with a lowercase *p*) involves coercive norm or rule enforcement. The implication is that teachers who enforce codes of conduct police; people who monitor parking meters and give out tickets police; and child protective services agents who tell parents to change their behaviors police. This definition also has the rather counterintuitive implication that abolitionist protesters may themselves be "police" if they take on distinctive policing roles like directing traffic.[32]

31 This issue was helpfully raised in detailed comments from an anonymous reviewer. Similar issues are discussed in Hunt, "The Limits of Reallocative and Algorithmic Policing, 29; and Monaghan, *Just Policing*. Monaghan in particular worries that the alternatives proposed by abolitionists and others still amount to police, as he defines policing broadly as coercive social control. As he claims, "all abolitionist alternatives involve social control and the policing of public space," and as such, "the risk of unjust policing is therefore always with us" (17).

32 Del Pozo, *The Police and the State*, 87–88.

The unbundling proposal implicates the first sense of Policing, maintaining that Police departments ought to have fewer roles and responsibilities than they currently do. Adopting the formal, institutional definition of Policing is required to render the familiar components of the proposal intelligible—e.g., "We shouldn't have Police enforcing traffic regulations." Correspondingly, if we employ the informal definition of policing, we cannot intelligibly claim that we should reallocate "police powers and responsibilities," given that policing, informally defined, is capacious and not confined to a discrete set of actors; we cannot reallocate a social practice. Moreover, the various actors who take up the roles currently assigned to Police will "police" in the broad, informal sense. Jake Monaghan makes this point, claiming that unbundling-style alternatives constitute policing as they are "still an expression of a claim of social control."[33]

The unbundling proposal is consistent with and even fundamentally based on the scholarship on policing (discussed above) that shows that Police do far more than what is specified by their formal, institutional duties; crime fighting and law enforcement comprise only a small part of the job. Proponents of unbundling agree with the important *descriptive* claim that the Police currently fulfill these roles beyond their narrow official mandate, but they resist the *normative* claim that this ought to be the case, arguing instead that we should carve off some of the roles and responsibilities currently aggregated within the social peacekeeper model of policing.

This normative claim seems to be in tension with some existing accounts of the distinctive normative capacities of policing. For instance, Brandon del Pozo argues that policing involves three central normative powers: impartial protection and rescue; arrest for the purposes of adjudication; and brokering and enforcing social cooperation.[34] Eric Miller argues that the police are "the agency authorized to act upon the state's duty to govern in response to public emergencies," and so one of their core powers is making "authoritative determinations about how to respond" to threats to public order.[35] What del Pozo's and Miller's pictures of policing share is that they extend the normative mandate of

33 Monaghan, *Just Policing*, 22.

34 Del Pozo, *The Police and the State*.

35 Miller, "The Concept of the Police," 573. Importantly, Miller also alludes to the appeal of proposals like unbundling, suggesting that "where other public officials or social institutions or individuals are better able to enforce the law, then the police should defer to those officials, institutions, or individuals and use their authority to support these others in their efforts to govern effectively. To the extent that the state has the resources to create differently skilled specialized agencies capable of deploying non-violent responses, the state itself fails in its governance obligations if it tasks organizations that are primarily trained and outfitted to respond with violence to fulfill these roles" (575). The latter point in particular speaks to my argument about minimally violent capacities in section 2.1 below. Though his goal is

policing well beyond its law enforcement function. However, both accounts begin with the Bittner-esque descriptive claim about police as generalist first responders and then proceed to ask what unifies (or constitutes "the political essence") of police work in light of this descriptive fact.[36] In other words, their analyses consider the catchall order-maintenance function of Policing to be fixed and then proceed to ask the normative question of how that function should be justified within a political system. As unbundling involves *reimagining* the current role of Policing, it is no surprise that these normative accounts are ill fitting.

What is the value of theorizing about the role of the Police in this formal, institutional sense? One of the oldest questions in political philosophy is how best to structure states and their social and political institutions. We are concerned about the scope of roles and responsibilities of many diverse actors within the criminal legal system, including prosecutors, judges, detectives, probation officers, and Police. Thinking about the limits of Policing roles is especially important given, as Monaghan helpfully describes, that policing is integrated into a complex and "coupled" criminal legal system with multilevel unfairness and failures.[37] In a system where "legislatures pass unjust laws, the trial system dolls out too much punishment, and background injustices make it more likely that certain groups get caught up in the criminal legal system," the actions of Police, like choosing to make an arrest, can have deeply concerning effects downstream.[38] One way to confront the problem of coupling is to question and reimagine formal Policing roles and responsibilities, in particular, to mitigate the extent to which order maintenance is coupled with the rest of the criminal legal system.

2. UNBUNDLING AND PRINCIPLES OF INSTITUTIONAL DESIGN

I will now attempt to convince you that we should unbundle policing by working through a series of arguments. Crucially, the case for unbundling is strongest if we take the arguments in tandem.

---

defining and explaining the core normative powers of policing as it stands, Miller views the current role of policing as a contingent social arrangement.

36 Del Pozo, *The Police and the State*, 10–25; and Miller, "The Concept of the Police," 580. See also Bittner, "The Police on Skid Row."

37 Monaghan, *Just Policing*, ch. 2.

38 Monaghan, *Just Policing*, 27.

2.1. Unbundling and Minimizing Violent Capacities

We all agree that police should use force only as a last resort, exhausting all nonviolent means of intervention first. In general, morality demands minimal use of state-sanctioned violence, reserving force against citizens only when all other options are exhausted. An analogous principle applies at another normative level: that of institutional design. In particular, we should distribute roles, responsibilities, and mandates between institutions so that institutions have violent capacities only as a last resort. Correspondingly, institutions endowed with violent capacities should have narrow mandates that are distinctly focused on interventions that require force, thus limiting their intervention in affairs that do not require violence.

This proposal already has broad institutional and public support: nobody wants SWAT teams to conduct health and safety inspections or to collect taxes. The military should not be tasked with enforcing school conduct codes. Why is this? First, echoing principles about self-defense, institutions should have violent capacities only when *necessary* to fulfill their ends: SWAT teams are not needed for health and safety inspections. Furthermore, the degree and capacity of violence should be *proportional* to the degree of harm (or risk of harm) it aims to prevent. If only armed military officials could effectively enforce school conduct codes, deploying them may still be impermissible given the menial harm of violating conduct codes.[39]

But why are unnecessary and disproportionate violent *capacities* objectionable? One reason is that the capacity for violence heightens the risk of harmful and deadly outcomes in the context of inevitable human error; thus, it is morally imperative that we distribute institutional labor to reduce the risk of these outcomes. Mistakes are more critical when they involve a gun as opposed to a baton or, better yet, a notebook for issuing citations. If we want to minimize the risk of unjustified deployments of violence and harm, we should minimize the scope and impact of institutions with violent capacities. In the case of policing,

39 An anonymous reviewer has rightly pointed out that I have not specified the types of cases where the capacity for violence is unnecessary or disproportionate; instead, I have focused on relatively uncontroversial examples to argue for unbundling. The reviewer has also noted that context matters; for example, traffic policing may require more forceful interventions if armed groups frequently take over intersections. In countries with widespread gun ownership, like the United States, more conflicts unfortunately demand interventions from violence responders (SVRs). Reducing gun prevalence is crucial because guns escalate the need for state-sanctioned violence. Additionally, rural areas may require higher per capita rates of SVRs due to the need for prompt emergency response.

reducing the scope and mandate of people wielding lethal force would reduce the number of people who are unjustifiably harmed by it.[40]

Indeed, police killings often occur in contexts where violent capacities are unnecessary and/or disproportionate. Traffic stops offer a clear example: Philando Castile, a thirty-two-year-old school cafeteria worker, was killed during a routine traffic stop in the Minneapolis–Saint Paul area in 2016.[41] His girlfriend and her four-year-old daughter were present. These sorts of cases are far too common: in the United States, police killed nearly six hundred people during traffic stops from 2017 to 2022.[42] Getting armed officials out of the business of traffic enforcement and other forms of routine order maintenance offers a straightforward path to curtailing violent escalations.

The broader issue is that the police are the repositories of coercive violence, but many order maintenance roles do not require violent capacities. The narrow class of situations that do require (potentially) violent interventions could be tasked to a specialized class of violence responders (SVRs) who are more narrowly equipped and training to intervene effectively. Furthermore, we should be attentive to how the intellectual, technological, and cultural baggage of one mandate impacts the capacity of institutions to effectively carry out other mandates; violent capacities engender norms and dispositions that undercut the ability of police to safely and effectively perform routine order maintenance.

2.2. Unbundling and Distributing Epistemic Labor

Many instances of police incompetence and misconduct stem from their lack of expertise. Cops are not trained to be social workers, conflict mediators, mental health crisis interventionists, homelessness outreach service providers, or school counselors, though they are tasked with responding to the complex social problems that fall within the purview of these areas of expertise. It is understandable that allocating this bundle of responsibilities—and the epistemic burdens that accompany them—to the police yields disastrous results.

40 This concern can also be framed in terms of risks to police legitimacy, meaning the risk that the political power of the police will be exercised improperly. According to Monaghan, the more burdensome a police tactic, the greater the risk of its illegitimate use, as more severe tactics require stronger justification. Reducing the use of violent tactics lowers the legitimacy risk. Monaghan argues that criminal patrols, which combine law enforcement and order maintenance, pose high legitimacy risk, and he recommends separating these functions within police agencies to create two distinct divisions. See Monaghan, *Just Policing* chs. 3–5, 8. I agree with the thrust of this proposal but argue that we should go a step further by removing police from the business of order maintenance all together.

41 See Cooper, "Philando Castile Shooting (2016)."

42 Levin, "US Police Have Killed Nearly 600 People in Traffic Stops Since 2017, Data Shows."

The social peacekeeping model is epistemically overdemanding, and there are instrumental epistemic benefits to the unbundling proposal.

Let us step back from the case of policing to think about this issue more broadly. In designing political institutions, we should allocate roles and responsibilities to those with the most relevant expertise and be sensitive to the relative epistemic burdens of roles. Designing educational curricula should be left to those trained to do so, and we should not require experts in pedagogy to mint currency or engineer transportation systems. Expecting them to fulfill these additional roles (without adequate training) would be epistemically overburdensome and would predictably result in shoddy currency or transportation systems. Moreover, training them to do *all* of these tasks is not a feasible solution, as the aggregate burden for knowledge and training is too high. In general, narrower and rigorously defined roles limit relative epistemic burdens, foster expertise, and drive better outcomes. The same principles apply to policing. We should not train police to be experts in all arenas of social peacekeeping but rather reallocate many of these responsibilities and their corresponding epistemic burdens to those with narrower and more relevant expertise. Friedman makes this point in defending an unbundling-esque proposal, arguing that "no single human being" can be at the same time a "forceful crime-fighter, empathetic interviewer and assistant of victims, collaborator with communities and social service agencies, [and] solver of crimes."[43]

The problem is not only that police lack relevant knowledge to fulfill social peacekeeping roles but also that the expertise, tactics, and norms distinctive to crime fighting are inappropriate and even dangerous in other domains. Framed otherwise, policing expertise interacts with and often undercuts the knowledge and dispositions required for other forms of order maintenance. The consequences of spillover from the crime fighter role to other functions of policing are especially pronounced in the case of mental health crises. Police are trained to assert control through commanding voices, intimidating postures, and readiness to use force. When people do not comply, officers escalate by closing in and raising their voices. These tactics directly conflict with the principles of effective mental health crisis intervention used by social workers: de-escalation through calm communication, maintaining physical distance, and fostering a sense of safety rather than control. Using forceful approaches in such situations often exacerbates the crisis; unsurprisingly, an alarming number of people experiencing serious mental illness have been injured or killed by police.[44] The

43 Friedman, "Disaggregating the Police Function," 981.

44 See Fuller et al., *Overlooked in the Undercounted*.

fact that many mentally ill people have been directly or indirectly involved in traumatizing encounters with the police only exacerbates this tension.

To summarize: much of the work currently assigned to police requires expertise and training that is well beyond the purview of the core functions of policing. Instead of increasing police officers' epistemic burdens, we ought to allocate many social peacekeeping responsibilities to those who are better trained, equipped, and experienced.

2.3. The Downsides of Consolidation

It is worth considering the potential downsides of consolidating violent capacities: If there were an institution specifically focused on violent interventions, would it have a problematic institutional ethos? Broadly, when an institution is given tools *x* and *y*, the members of the institution tend to become especially invested in *x* and *y* and may come to think that *x* and *y* are the solution to everything. (If you have a hammer, everything looks like a nail.) Would consolidating violent capacities and violence-relevant expertise within a single institution lead to an overzealous deployment of violence—the very thing that the unbundling proposal is meant to avoid in the first place? These are serious concerns and, ironically, reflect the critiques of currently bundled police institutions. Narrowly focusing the mandate of SVRs to potentially violent situations addresses some of these issues. As opposed to status quo order-maintenance policing, SVRs would be deployed only when the threat of violent escalation is high. We may still worry about the conduct of SVRs when they *are* deployed. Human error will inevitably result in SVRs being sent to situations where violent interventions are not strictly necessary.

What is a viable solution, then? In my view, this problem is ripe for the familiar use-of-force and training reforms proposed in philosophy and elsewhere. Even if violent situations comprise the mandate of SVRs, official training and procedures should emphasize minimalism in use of force, aiming first to resolve violent situations without gunfire. Official policy should require officers to encourage armed and dangerous people to surrender peacefully and, when force is necessary, to use low levels of physical force or nonlethal weapons (Tasers, etc.). These changes to use-of-force policies and practices should be accompanied by oversight tactics—for instance, requiring SVRs to wear body cameras, otherwise holding bad actors accountable, etc. We should also be attentive to the psychological profiles of candidates when making SVR hiring decisions and disqualify those who are predisposed to use violence. One upshot of implementing these widely proposed reforms in the context of unbundling is that they can be more narrowly focused on a smaller and specialized group of actors and thus more feasibly achieved than attempting to reform policing writ large.

3. INJUSTICE AND UNBUNDLING

Unlike the preceding arguments, which appealed to general principles of institutional design, I will now defend an approach to unbundling that focuses on how policing institutions function in unjust and unequal social contexts. The notion that we ought to restructure policing because the institution is inherently racist and/or unfair is familiar in popular discourse and underwrites the organization of social movements. My goal in this section is to lay out a path from (1) policing's unjust functions to (2) the unbundling proposal. Though the sentiments that motivate the argument stem from critical policing activism, its structure is novel and applies to other justice-undermining institutions. Let us start by specifying the justice-undermining effects of policing institutions. In my view, injustice in policing results from the interaction of the following phenomena.[45]

> *Disadvantage Selection*: People who are marginalized around socially salient dimensions (race, class, addiction status, etc.) are disproportionately likely to have encounters with police.[46]
>
> *Policing Harm*: Encounters with police are (1) harmful, (2) risk harmful escalation, and/or (3) involve people in a harm-causing criminal legal system.[47]
>
> *Explicit and Implicit Bias*: Police have deep-rooted implicit and explicit biases against members of marginalized groups, which heightens the risk of harmful interactions between police and members of those groups.[48]

45 The outcome-focused account presented here differs from the view that racism in policing is rooted in racist attitudes or beliefs. It is also more capacious than Joseph Heath's suggestion that the problem with race and policing stems from how minority groups are "subject to coercive enforcement of social norms and standards of respectability that reflect parochial aspects of majority culture" ("The Challenge of Policing Minorities in a Liberal Society," 3).

46 There is a vast body of empirical work supporting this. For recent and historical data, see the databases of Mapping Police Violence and the *Washington Post*. Books on the topic include Davis, *Policing the Black Man*; Butler, *Chokehold*; and Zimring, *When Police Kill*. Note that gender is an exception to disadvantage selection, as women, though marginalized, tend to be arrested and incarcerated at lower rates than men.

47 That policing encounters are harmful or risk harmful escalation is supported by data on lethal and nonlethal police violence. See the databases of Mapping Police Violence and the *Washington Post*. On the harms of involvement with the criminal legal system, see Western, "The Impact of Incarceration of Wage Mobility and Inequality"; and Kirk and Wakefield, "Collateral Consequences of Incarceration."

48 An extensive body of data shows that police (as well as the general public) are more likely to view Black men as dangerous and act on those biases. See, e.g., Correll et al., "The Police

Disadvantage selection manifests in context-specific ways and results from a variety of factors.[49] Patterns of criminalization—such as criminalizing behaviors linked to addiction and homelessness—mean that membership in some marginalized groups overlaps with the content of the criminal law. Social strain, a result of systemic injustice, also leads marginalized people to engage in some criminalized behaviors at higher rates.[50] Additionally, even when there are similar rates for offenses across groups, marginalized people are disproportionately stopped, searched, and arrested due in part to the concentration of police in minority neighborhoods and biased assumptions about guilt and dangerousness. In the United States and elsewhere, this process is clearly racialized, with Black people overrepresented across all categories of policing interactions from routine stops to violence escalations, both now and historically.[51]

The interaction of disadvantage selection and policing harm means that policing disproportionately burdens those who are already marginalized, thus perpetuating and exacerbating endemic inequalities. Stated plainly, if members of marginalized groups are more likely to have encounters with police, and those interactions are (1) harmful, (2) risk harmful escalation, or (3) involve people in a harmful criminal legal system, then members of marginalized groups will be disproportionately harmed by policing. The risk of harm increases for those against whom police hold implicit or explicit biases.

Such entrenched endemic injustice means that marginalized people bear the burdens of harmful police interactions and violent escalations. Policing plausibly operates as a form of structural injustice, which occurs "when social processes put large groups of persons under systematic threat of domination or deprivation of the means to develop and exercise their capacities," as Iris Marion Young describes.[52] Like other forms of structural injustice, injustice in policing is not reducible to intentional actions but arises from the very constitution of our social structure, from the interface between background conditions and core features of our existing institutions. This aspect of policing may

Officer's Dilemma."

49 Context determines what categories are socially salient. Some groups that may be overrepresented in police interactions include religious and linguistic minorities, racial minorities, people with disabilities, people with HIV, recent immigrants, queer people, Indigenous people, and people who are low income and poor (among others).

50 For instance, people are more likely to participate in the illicit economy if their employment prospects are limited. For an overview of strain theory, see Agnew, *Pressured into Crime*.

51 For a helpful overview on the evidence of racial disparities in police practices, see Ghandnoosh and Barry, "One in Five."

52 Young, *Responsibility for Justice*, 54.

go unnoticed if we adopt certain ideal theoretical assumptions—for instance, if we theorize on the basis of equal background conditions or assume that social institutions function only as intended.[53] By rooting our analysis in the realities of our unequal and complex social world, we subvert this normative picture in favor of a more complex but pragmatic understanding of policing and its impact on social structure.

At the beginning of *A Theory of Justice*, Rawls claims that "laws and institutions, no matter how well-arranged and efficient they are, must be reformed or abolished if they are unjust."[54] Engaging with Rawls on policing, Jake Monaghan rightly argues that police agencies are entitled to the status of major institutions in the basic structure of society.[55] As such, we should consider the extent to which they promote or undermine social justice and, in particular, how they impact the well-being of the most marginalized.[56] If policing—or any major social institution—has equality-undermining effects, we should do something about it.

Determining how to mitigate injustice in institutions raises highly context-sensitive questions at the intersection of philosophy and public policy. Broadly, we can specify two varieties of remedies to institutional injustice. The first variety is *intra-institutional*. Intra-institutional measures address injustice by installing changes *within* existing institutions while keeping their basic structure and mandate in place. Some familiar intra-institutional remedies include changes to formal rules and policy, training, the distribution of labor, and decision-making procedures, as well as the elimination of bad actors. These strategies may be accompanied by others focused on changing informal norms

53 I am thinking of Charles Mills's helpful schema of problematic idealizations in Mills, "Ideal Theory as Ideology."

54 Rawls, *A Theory of Justice*, 3.

55 Monaghan, "Idealizations and Ideal Policing."

56 Rawls focuses on the justice of the basic structure, which is constituted by many social institutions. My approach differs because I focus on a specific institution (policing) and ask if it tends to exacerbate injustice. One may wonder whether policing, assessed alone, compounds injustice but is part of a broader system that is just, or if police would have any role to play in an ideal society. (For discussion of these issues, see Fryer, "Idealizing Abolition"; and Monaghan, "Idealizations and Ideal Policing.") I do not address either of these questions here as my interest is in the more focused, nonidealized issue of how we can move towards a *rough* outline of justice on the basis of reasonable and noncontroversial assumptions—in other words, which changes would result in a more fair and equal world (holding fixed some considerations about background injustice, human psychology, and so on). I thank an anonymous reviewer for helping me to refine this methodological orientation.

and culture.[57] The second family of remedies is *extra-institutional*, meaning that they attempt to combat institutional injustice by reimagining the structure of an institution as it relates to others within a wider institutional system. Some familiar measures of this variety include oversight by other organizations, a redistribution of institutional powers and responsibilities, and the abolition of institutions all together.

I now advance two normative principles about these forms of remedy. First, when intra-institutional measures fail to mitigate injustice in institution *X*, we ought to experiment with extra-institutional measures to reallocate the power, roles, and responsibilities assigned to *X* (the *experimentation requirement*). Second, if an alternative institution (or set of institutions) *A* can adequately fulfill the roles and responsibilities currently assigned to *X* and avoid the justice-undermining effects of *X*, we should reallocate roles and responsibilities to *A* (the *reallocation requirement*). Both principles apply to policing and ground the injustice-based case for unbundling.

The rationale for the experimentation requirement is straightforward: if we have two paths to mitigate a pressing problem, and one proves inadequate, we should explore the other. In this case, when piecemeal reforms fail to combat injustice, we should experiment with broader transformative solutions.[58] Experimentation sometimes involves creating *new* institutional forms aimed at fulfilling socially valuable functions. For example, if our aim is to ensure that everyone has a basic minimum level of economic security, and we find that means-tested welfare systems often fail to reach those who need assistance due to administrative inefficiency, we may turn to experimenting with an entirely new system to foster the basic minimum (a universal basic income scheme, for instance).

Let us apply the experimentation requirement to policing. The first question to consider is whether the problems with policing are immune to incremental reform, as activists often claim.[59] As framed here, the meat of the critique is that

57 Addressing culture is important given how in cases of structural injustice, "informal social norms and institutional rules generally work in tandem. They continuously interact, each begetting, reshaping, sustaining, or undermining the other. . . . Adherence to informal norms sometimes continues after formal rules governing the relevant conduct are abandoned" (Powers and Faden, *Structural Injustice*, 100).

58 We should attempt piecemeal, intra-institutional reforms first because of (1) inevitable transition costs and (2) feasibility concerns associated with broader inter-institutional measures. While this may provoke concerns of status quo bias, the thrust of the argument still applies if we hold this critique: we may say we should try both intra- and inter-institutional measures (versus prioritizing one over the other), but if intra-institutional measures fail, we should focus on inter-institutional ones.

59 Kaba, "Yes, We Mean Literally Abolish the Police"; and Vitale, *The End of Policing*.

we cannot eliminate injustice in policing through intra-institutional measures. To support this, critics often appeal to how substantial reforms have been implemented in places where egregious policing practices continue. The city of Atlanta, for instance, implemented a range of police reforms over decades; despite these reforms, police-initiated killings continued.[60] However, it is difficult to draw a general conclusion from this piecemeal evidence, especially because there is such a broad range of intra-institutional reforms to policing on the table: examples include changing the laws and policies that govern police tactics, holding bad actors accountable, rethinking and expanding police training, and so on. To claim that all of these efforts fall short in mitigating injustice is a difficult empirical and philosophical project.

A more persuasive way of showing that policing is uniquely immune to reform appeals to how injustice is sustained in policing. The problem with intra-institutional reforms is that they do not address the interaction between disadvantage selection, policing harm, and implicit and explicit bias identified above. More specifically, solutions internal to policing do not undercut disadvantage selection or explicit and implicit biases, since they result from broader features of the social structure that are well beyond the scope of policing behaviors and practices.[61] Reforms to policing do not mitigate the inequality and social strain that drive disparities in policing, nor the endemic biases that sustain it. Moreover, reforms aimed at curtailing the harms of policing are always limited, given the ways that police are (in the formal, institutional sense here) repositories for state-sanctioned violence. Simply eliminating their capacity to harm runs contrary to their design function and fundamental social role as a coercive institution in the context of a criminal punishment system that *aims* to harm those who break the law.[62] Without disrupting police culture or changing its basic technologies, structure, and institutional mandate, we cannot hope to prevent or mitigate the harm that police cause to the people they interact with.

Because of this underlying mechanism, injustice in policing is immune to intra-institutional measures, and so we should experiment with extra-institutional means, per the experimentation requirement. In particular, we should try

60 Herskind and Roberts, "The Failure of Police Reform."

61 One may hope that implicit and explicit bias training could remediate biased behaviors, but evidence of this is lacking. See Lai and Lisnek, "The Impact of Implicit-Bias-Oriented Diversity Training on Police Officers' Beliefs, Motivations, and Actions." Moreover, eliminating biases would not disrupt the interface between disadvantage selection and policing harm that sustains injustice.

62 Furthermore, if we were to "change" policing so as to be noncoercive, decoupling it from use of force and the broader criminal legal system, it would cease to be policing (in the formal, institutional sense).

to create new institutional forms aimed at fulfilling the socially valuable functions of policing. Unseating the status quo and envisioning radical alternatives in the way the requirement suggests are key tenets of abolitionist social thought.

Beyond experimenting with transformative alternatives, the reallocation requirement provides a positive case for unbundling now. Recall that the requirement states that if an alternative institution (or set of institutions) *A* can adequately fulfill the roles and responsibilities currently assigned to *X* yet avoid the justice-undermining effects in *X*, we should reallocate roles and responsibilities to *A*.[63] The requirement is based on a simple Pareto superiority-style principle: if there are two potential institutional design proposals (that both offer to fulfill some valuable functions), and one has a flaw that the other avoids, we should install the latter.

In order for the principle to motivate unbundling, two premises must hold true. The first is that alternative institutions *can* adequately fulfill some of the socially valuable functions at which policing aims; that reallocation is theoretically and practically possible.[64] This condition is met in the case of unbundling because many of the roles and responsibilities that unbundlers propose for reallocation are already the business of other institutions and tangential to the distinctive crime-fighting mandate of police. On one hand, we can envision expanding the mandates of existing institutions; on the other, we can draw on models implemented elsewhere to create new ones. The second premise holds that unbundling reduces the justice-undermining effects of policing—i.e., by unbundling, we mitigate injustice.[65] This premise is more contentious but plausible given the injustice-sustaining mechanism identified. Because unbundling

63 The principle holds if you substitute other negative attributes for justice-undermining effects. For instance, if an alternative institution (or set of institutions) *A* can adequately fulfill the roles and responsibilities currently assigned to *X* yet avoid the disutility in *X*, we should reallocate roles and responsibilities to *A*.

64 Reallocating the powers of arrest from police to other institutions (for instance) is not *theoretically* possible because the new institutions would still be police (by the formal, institutional definition employed here). Relatedly, we may say that reallocating criminal investigation from police to other institutions is not *practically* possible given that investigations require access to confidential data, forensic analysis, and coordination with courts—expertise and legal authority that civilian agencies simply do not possess. However, practical limitations are often the product of laws and extant institutional structures that we may also aim to change.

65 As I discuss with respect to the replication problem below, alternative nonpolice responses may still be harmful, albeit to a significantly lesser extent. An anonymous reviewer has helpfully pointed out that because of this, the constraint amounts to a defense of *minimizing* injustice-promoting harms versus eliminating them altogether. So the "avoid the unjust effects" clause of the constraint can be read in terms of avoiding *the degree of* unjust effects (i.e., significantly reducing them).

separates violent capacities and criminal punishment from social peacekeeping, it undercuts policing harm. This is because if we unbundle, fewer social problems would be handled by those authorized to use force. Instead, the governing norms of intervention would be care and support, thus limiting the *immediate* harm of police interactions. Unbundling disrupts the longer-term harms of policing by breaking the link between social peacekeeping and the criminal legal system's downstream effects.

The overarching point is that if we are serious about curtailing the unjust effects of policing, we ought to reduce the size and scope of policing institutions altogether. A metaphor illustrates the thrust of the arguments here: when a machine is broken (in the sense that fulfilling its intended aims has worrisome, unwelcome consequences), we ought to consider creating another machine. If there are other machines to which we can reallocate some of the broken machine's jobs and avoid those consequences, we should do that in the meantime.

4. Objections to Unbundling

I will now work through three lines of objection to unbundling. The objections reflect concerns raised in divisive public debates about policing and, in particular, responses to the ideas that we should defund or abolish the police—demands closely related to unbundling, as previously discussed.

4.1. Will Unbundling Lead to Increased Crime?

Some skeptics of unbundling claim that policing is a necessary evil. While they acknowledge the many problems with our current policing system, they believe that police (in their order-maintenance role) are essential to keep crime at a manageable level. A similar argument arises in discussions about jails and prisons, with critics fearing that decreased incarceration would lead to unacceptably high crime rates.[66] There is also an egalitarian concern in this view, as higher crime rates may disproportionately harm marginalized people because they are more likely to be victims.[67]

This objection is a reasonable response to some understandings of police abolitionism; if there were *no police*, potential offenders may not be deterred from engaging in crime (knowing they would not be caught), and so crime would drastically increase. As I am not arguing for abolitionism here, the

66 The issues of necessity, crime prevention, and prison abolitionism are discussed at length in Boonin, *The Problem of Punishment*, ch. 5.

67 The disproportionately burdensome effects of increased crime must be balanced with the disproportionately beneficial effects of reduced police violence.

objection to my proposals is more nuanced: if we were to reallocate many order-maintenance powers and responsibilities from police to other institutions, crime rates would reach an unsustainable level—and thus, we should not unbundle policing. I think this claim is empirically and normatively suspect.

One may think that because unbundling reduces the size, scope, and impact of policing, there will be fewer police and thus more crime; however, this claim constitutes a vast oversimplification of the extensive literature on crime deterrence and policing. What police are *doing* matters immensely. Studies of police deterrence often evaluate the relative efficacy of three sorts of police practices: random patrol, rapid response, and reactive investigation. The deterrent effects of random patrol are most relevant to evaluating unbundling.[68] Reviewing the literature on police deterrence, Daniel Nagin, Robert Solow, and Cynthia Lum claim that "there are good reasons for skepticism about the efficiency and effectiveness of random patrol."[69] In addition to the lack of evidence indicating a negative correlation between crime rates and the number of officers on patrol, they appeal to the psychology of deterrence, claiming that police deter crime primarily by reducing the perception of would-be offenders that crimes can be committed successfully without them being apprehended.[70] Given that it is unlikely that police on patrol would be in the right place at the right time so as to increase a would-be offender's assessment of risk, randomness is significantly less effective than targeted strategies like hot-spot policing and "focused deterrence."[71] Thus, the idea that fewer cops on the beat engaging in order-maintenance activities leads to more crime is empirically suspect.

68 The role of police in criminal investigations raises important questions. Currently, investigations are often ineffective—especially in cases of gender-based and sexual violence—and employ controversial interrogation tactics (Venema, "Police Officer Schema of Sexual Assault Reports"; Du Mont et al., "The Role of 'Real Rape' and 'Real Victim' Stereotypes in the Police Reporting Practices of Sexually Assaulted Women"; and Hunt, *Police Deception and Dishonesty*). Police departments and organizations also frequently fail to adequately respond to officer misconduct (Armacost, "Organizational Culture and Police Misconduct"). This has led to the establishment of alternative investigative bodies like the Civilian Complaint Review Board in New York City. We might consider reallocating the responsibility of crime investigation from police to an independent specialized entity because of these concerns, though doing so would require significant changes to existing laws and policies.

69 Nagin et al., "Deterrence, Criminal Opportunities, and Police," 77.

70 Nagin et al., "Deterrence, Criminal Opportunities, and Police," 77–78.

71 See Braga et al., "Hot Spots Policing and Crime Reduction." Thinking about hot-spot policing and focused deterrence through the lens of unbundling raises a number of interesting issues. A major concern with hot-spot policing is that it disproportionately targets low-income communities of color, which can lead to overpolicing, police violence, and an increased sense of surveillance, ultimately eroding police-community relations. Hot-spot

The assumption that fewer police leads to more crime also overlooks the relevance of confounding variables. Decades of criminology and sociology show that crime rates are shaped by social, economic, and environmental factors beyond policing. This complicates the calculation of the deterrent effect of increased police presence but also highlights other ways to reduce crime. For example, crime rates tend to decline with the rise of community nonprofits.[72] Since crime often stems from a lack of resources, investing in support and care upstream significantly impacts crime rates. Unbundling is not only a negative proposal aimed at dismantling police but also a proactive approach to crime prevention, including, in particular, installing targeted interventions for mental health, homelessness, and addiction and supporting community-led violence prevention efforts. Thus, in evaluating whether unbundling would lead to increased crime, we should consider not only the effects of decreased police presence but also the preventive effects of its positive vision.

Finally, even if unbundling were to lead to increased crime, we need not dismiss the proposal on that basis. This is because normatively, our goal is not to maximize crime reduction but rather to balance it with other morally valuable ends. We could ensure that there was very little crime by forcing everyone to stay in their homes or by recording their every move. We could preemptively incarcerate people without due process. Despite their potential efficacy for crime reduction, these strategies undercut our rights at grave moral costs, and so installing them is impermissible. Analogously, if crime were to increase as a result of unbundling, this is not sufficient ground for dismissing the proposal; rather, we need to balance the (unclear) degree of crime prevention with other ends. Given the gravity of the problems with catchall order-maintenance policing described, unbundling may still be the right path forward, despite leading to some level of increased crime.

4.2. Will Unbundled Institutions Replicate the Problems with Policing?

Another objection to unbundling is that it will not solve the problems with policing that motivate us to restructure the institution in the first place because

policing and similar strategies have also proven to be effective in reducing violent crime. It is important to note, however, that some of the most effective hot-spot policing initiatives, such as Chicago's Ceasefire and Cure Violence, leverage community violence interruption by employing community members as street-level interventionists. Interventionists mediate conflicts, provide cash assistance, and support victims in hospital settings. These efforts exemplify the type of trained civilian de-escalation advocated here. At the same time, they are frequently accompanied by police presence and the looming threat of severe sanctions to deter retaliation. Is crime reduced because of community violence interruption or by greater police presence?

72 Sharkey, "Community and the Crime Decline."

the alternative, nonpolice institutions (that will take on the roles and responsibilities currently allocated to the police) will be subject to similar critiques.[73] As described by Monaghan, order maintenance always involves policing (in the broad, informal sense), and "the risk of unjust policing is therefore always with us."[74] This issue often arises within the police abolitionist movement. Mariame Kaba and Andrea J. Ritchie warn us of "the authority figures who make up the 'soft police'—including medical professionals, social workers, and government bureaucrats" who "engage in policing in their own right, and are often entangled with traditional law enforcement."[75] Correspondingly, they critique slogans such as "counselors not cops," "caseworkers not cops," and "treatment not punishment" because a police-free world should not involve the one-to-one replacement of police with other coercive public institutions.

Many existing social service institutions are fundamentally flawed; for instance, systems of mental health care routinely confine people against their will in brutal institutions with unfair and nontransparent procedures, often without access to legal counsel.[76] Liat Ben-Moshe characterizes psychiatric hospitals as medicalized carceral spaces and argues that they ought to be abolished.[77] Drug treatment institutions are subject to similar critiques.[78] Bernardo Zacka points to one cause of these failures, arguing that bureaucrats involved in the provision of public services, like social workers, teachers, and police officers, operate in adverse institutional conditions that tend to erode their moral sensibilities and "truncate their understanding of their role and responsibilities."[79] These concerns caution us against silver-bullet thinking about alternatives to police.

I agree with Kaba and Ritchie's ultimate contention that we need to respond to the replication critique with careful institutional design that is cognizant of how nonpolice institutions can engage in policing-adjacent practices that give

73 This objection is closely related to the question of whether unbundling will reduce injustice in policing (discussed in section 3 above).

74 Monaghan, *Just Policing*, 17.

75 Kaba and Ritchie, "No More Police." Note that these agencies do not count as police in the formal, institutional sense.

76 Objectionably, in many states, there is not a right to counsel in cases of family law, involuntary commitment, and medical treatment. See Brito, "The Right to Civil Counsel."

77 Ben-Moshe, *Decarcerating Disability*.

78 See McCorkel, "The Second Coming."

79 Zacka, *When the State Meets the Streets*. In particular, Zacka argues, the everyday demands of their work predispose public servants to adopt reductive dispositions (specifically those of indifference, enforcement, and caregiving) that cause them to lose touch with the plurality of demands relevant to moral decision-making. However, Zacka argues, this is an understandable response to the psychological pressures of the direct public service in which they are engaged.

rise to the problems canvassed before. I mentioned earlier that my goal is not to offer a complete package of institutions that should take up the powers and responsibilities currently tasked to police, so responding to the charge that the new institutions will replicate injustice is dialectically challenging. Nevertheless, in reimagining our institutional landscape, we should aim not only to *reform* our existing institutions but also to question and reconfigure their basic structure, being attentive to how they graft onto complex and unjust social landscapes.

But just because we need to be careful in crafting alternatives does not mean we should abandon the unbundling proposal altogether. Importantly, the unbundling proposal clearly assuages two of the central issues with policing: (1) unjustified use of violence and force and (2) inappropriate responses to social issues such as addiction, homelessness, and mental health crises. So it is a mistake to say that unbundling replicates the problems with policing in that the problems are either identical or equally grave. The deprivation of *benefits* resulting from injustice in social services is clearly objectionable; however, I think it would be a moral mistake to equate these harms with those of policing. Normatively, the unjust use of state-sanctioned violence is more objectionable than the deprivation of social benefits; more broadly, burdens and benefits are not on par in this way.[80]

Nonetheless, support-based interventions are at times harmful and perpetuate inequality. My pessimistic view is that in unequal societies, social problems sometimes require coercive treatment that inevitably risks harming the most vulnerable. What morality demands we do about this unfortunate fact is structure our institutions so as to minimize those unfair burdens. I will now sketch a few principles aimed at minimizing harm and the unfairness it engenders.

First, social service interventions should be *minimally coercive*, with restrictions on individual liberty—such as confinement, surveillance, and forced treatment—used only as a last resort in cases where people are imminently dangerous to themselves or others. The principles of proportionality and necessity that govern the use of force should also guide these more moderate interventions, and noncoercive, consensual forms of care and support should be the default. Moreover, we should implement procedural measures that foster *transparency* and *community governance*. Transparency requires well-defined

80 Various normative ideas underwrite this. On one hand, we may think there is something especially egregious about police-initiated violence. As Monaghan argues, police may have special moral obligations that make violence initiated by them worse than that initiated by private citizens ("The Special Moral Obligations of Law Enforcement"). Beyond this, there are important moral differences between harming and not aiding—and in this case, state-sanctioned violence is a harm, while state agencies failing to provide benefits constitutes failure to aid.

protocols for navigating systems of support. People should also have access to representatives to advocate for them as they work through public systems, such as legal counsel in cases of mandated treatment or confinement. Community governance structures empower those who are most impacted by social policies to have a stake in designing and administering them, and as a rule, we ought to prefer local, community-based responses to large state bureaucracies. There are a variety of ways to install self-governance structures. Some possibilities include sortition-selected decision-making bodies or citizen advisory committees composed of people with direct experience navigating public service systems.[81]

Finally, as Kaba and Ritchie argue, designing fair institutions requires that we question our ideas and, in particular, our tendencies to "continue to control currently criminalized people and populations by placing them 'Somewhere Else,'" which requires "Someone Else—if not police—to put them there."[82] The suggestion here is to critically examine exclusionary ways of thinking about social problems, challenging assumptions about what and who is considered "normal" and rethinking which situations warrant intervention at all. Monaghan makes a similar point in arguing that police should not intervene in many circumstances where there is a disagreement about what constitutes acceptable use of public space.[83] Working to address the upstream factors driving social disorder while also questioning our ideas of what is disruptive and normal will help prevent unbundling institutions from replicating the problems with our present policing regimes.

4.3. Will Unbundling Work in Places with High Rates of Gun Ownership and Gun Violence?

High rates of gun ownership, especially in the United States, may pose a challenge to the unbundling approach. In critiquing reallocative policing (a similar proposal), Luke Hunt refers to this as the "socio-scientific problem."[84] Hunt cites the facts that there are more civilian-owned firearms (393 million) than people (326 million) in the United States, and people choose to use these guns far too often (for example, in response to others playing music loudly in a car or failing to turn off their phone in a movie theater).[85] People regularly carry

81 For more on sortition and citizen advisors, see Guerrero, "Against Elections"; and Landemore, *Democratic Reason*.

82 Kaba and Ritchie, *No More Police*, 148.

83 Monaghan, *Just Policing*, ch. 6.

84 Hunt, "The Limits of Reallocative and Algorithmic Policing."

85 Hunt, "The Limits of Reallocative and Algorithmic Policing," 9.

firearms in public, and there is an alarmingly high incidence of mass shootings.[86] Do high rates of gun ownership undermine the appeal of unbundling?

Let us break down this objection more clearly. There are three possible concerns: (1) unbundling policing would lead to an increase in gun violence; (2) responses to gun violence would become inadequate; and/or (3) those taking on roles currently assigned to police—like social workers or traffic patrollers—would be at risk due to the prevalence of guns. The first concern about increased gun violence mirrors the crime increase objection already addressed above, so I will focus here on 2 and 3.

Let us begin with responding to gun violence. As Hunt rightly notes, police are "rarely spatiotemporally present at the scene of crime to stop assailants in the act"; rather, they are usually called to the scene.[87] Unfortunately, police responses to gun violence are often inadequate, particularly in the case of mass shootings. Critics of the response to the 2018 Parkland school shooting highlight how communication issues, coordination problems, and the lack of an immediate response plan led to preventable fatalities.[88] Effectively responding to these high-stakes crises requires training, expertise, and practice. A key component of unbundling is the establishment of a specialized class of violence responders who can more effectively intervene in dangerous situations—so plausibly, responses to gun violence would improve if we unbundled policing.

One important idea emanating from this objection is that the size and prevalence of SVR institutions should depend on contingent social factors, in particular, rates of gun ownership and gun violence. Rates of gun *violence* are particularly relevant for SVR allocation, as some places with high rates of gun ownership like Vermont and Switzerland nevertheless have low rates of gun violence.[89] A quick and efficient response is crucial for effectively addressing gun violence, which should inform decisions about the number of SVRs and their deployment. In rural areas, a higher ratio of SVRs per capita may be necessary to ensure timely responses.[90]

86 See the statistics at the Gun Violence Archive (GVA), https://www.gunviolencearchive.org.

87 Hunt, "The Limits of Reallocative and Algorithmic Policing," 9–10.

88 Thompson, "To Stop a Shooter."

89 See Stroebe et al., "Gun Ownership and Gun Violence"; World Population Review, "Gun Ownership by State"; and National Center for Health Statistics, "Firearm Mortality by State, 2022."

90 I mentioned above that in one view, unbundling carves off many of the powers and responsibilities of policing, reducing its size and scope overall but leaving police departments intact specifically to serve SVR functions. The alternative view is that we should create SVRs from the ground up given the objectionable features of police culture. Regardless of the view you adopt, we may go further to say that existing police department locations should

Let us turn to the concern about the safety of social workers, traffic patrollers, and mediators who will take on roles currently assigned to police. As a baseline, police are often unsafe—the fact that they have guns and powers to arrest does not always protect them. One of the many unfortunate facts about high rates of gun ownership is that people tasked with responding to crises can be victims of gun violence in the process. The worry with unbundling is that more people will be at higher risk if we restructure institutional responsibilities. It is true that to some degree, unbundling may involve shifting the risk of experiencing violent interactions from police officers to other actors. Indeed, street outreach conducted by social workers does involve some level of physical danger.[91] However, shifting the risk in this way is not necessarily bad. Moreover, to minimize risk, SVRs should accompany other responders in situations where there is known to be a high risk of violence. Furthermore, dispatching nonpolice actors may lead to lower rates of violent interactions, as they will not employ the coercive tactics that often escalate conflicts. For instance, traffic enforcers could avoid high-speed chases for minor infractions by opting instead for lower-risk approaches like sending tickets by mail. Tactics that minimize the risk of escalation are essential for protecting anyone involved in potentially dangerous situations.

The prevalence of guns contributes to a barrage of social problems and makes managing them more hazardous. Instead of assuming gun prevalence is fixed and designing our social institutions around it, we should also (obviously) focus on reducing the number of guns. Doing so would enable us to create fairer, safer, and more effective institutions.

5. CONCLUSION

I hope to have left the reader with a clear sense of what the unbundling proposal involves and to have distilled the most persuasive rationales for realizing it. Again, the case for unbundling is strongest when we consider the arguments for it in tandem. One broad insight emanating from this discussion is that the moral questions of policing extend well beyond the scope of individual interpersonal ethics, implicating broad political philosophical issues about what sorts of institutions we should install to deal with complex social problems in our messy and unequal world. The consistent failures of extant institutions encourage us to imagine alternative institutional forms. In jurisdictions around the

remain in place, and there should be far more SVRs than may plausibly be deployed at any time.

91 See Spencer and Munch, "Client Violence Toward Social Workers."

world, efforts to unbundle are currently underway; the experiment is happening! These efforts will offer further insight about how to promote public safety without police. Beyond the case of policing, reimagining our present criminal legal institutions is both necessary and urgent. Together, we can build safe communities and create/recreate institutions grounded in justice and compassion.[92]

University of California, Santa Cruz
lyons@ucsc.edu

REFERENCES

Agnew, Robert. *Pressured into Crime: An Overview of General Strain Theory*. Oxford University Press, 2006.

Armacost, Barbara E. "Organizational Culture and Police Misconduct." *George Washington Law Review* 72, no. 3 (2004): 453–546.

Asher, Jeff, and Ben Horowitz. "How Do the Police Actually Spend Their Time?" *New York Times*, June 19, 2020. https://www.nytimes.com/2020/06/19/upshot/unrest-police-time-violent-crime.html.

Balko, Radley. *Rise of the Warrior Cop: The Militarization of America's Police Forces*. PublicAffairs, 2013.

Bell, Monica C. "Police Reform and the Dismantling of Legal Estrangement." *Yale Law Journal* 126, no. 7 (2017): 2054–150.

Ben-Moshe, Liat. *Decarcerating Disability: Deinstitutionalization and Prison Abolition*. University of Minnesota Press, 2020.

Bittner, Egon. "Florence Nightingale in Pursuit of Willie Sutton: A Theory of the Police." In *The Potential for Reform of Criminal Justice*, edited by Herbert Jacob. Sage Criminal Justice System Annuals. Sage Publications, 1974.

———. "The Police on Skid Row: A Study of Peace Keeping." *American Sociological Review* 32, no. 5 (1967): 699–715.

Bondurant, Samuel R., Jason M. Lindo, and Isaac D. Swensen. "Substance Abuse Treatment Centers and Local Crime." *Journal of Urban Economics* 104 (2018): 124–33.

Boonin, David. *The Problem of Punishment*. Cambridge University Press, 2008.

Braga, Anthony A., Brandon S. Turchan, Andrew V. Papachristos, and David M. Hureau. "Hot Spots Policing and Crime Reduction: An Update of an

92 I am thankful for comments from Doug Husak, Alex Guerrero, John Oberdiek, Mike Otsuka, Gideon Rosen, Jennifer Page, two anonymous reviewers, and participants at the University of North Carolina Philosophy of Policing Workshop and the Rutgers University Dissertation Seminar.

Ongoing Systematic Review and Meta-Analysis." *Journal of Experimental Criminology* 15, no. 3 (2019): 289–311.

Brito, Tonya L. "The Right to Civil Counsel." *Daedalus* 148, no. 1 (2019): 56–63.

Brown, M. C., II and Camille Lloyd. "Black Americans Less Confident, Satisfied with Local Police." Gallup News, September 18, 2023. https://news.gallup.com/poll/511064/black-americans-less-confident-satisfied-local-police.aspx.

Butler, Paul. *Chokehold: Policing Black Men.* New Press, 2017.

Chyn, Eric. "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children." *American Economic Review* 108, no. 10 (2018): 3028–56.

Clifford, Stephanie. *Policing, Mental Illness, and the Media.* Palgrave Studies in Crime, Media, and Culture, 2021.

Cooper, Patrick. "Philando Castile Shooting (2016)." EBSCO Research Starters, 2021. https://www.ebsco.com/research-starters/law/philando-castile-shooting-2016.

Correll, Joshua, Bernadette Park, Charles M. Judd, and Bernd Wittenbrink. "The Police Officer's Dilemma: Using Ethnicity to Disambiguate Potentially Threatening Individuals." *Journal of Personality and Social Psychology* 83, no. 6 (2002): 1314–29.

Davis, Angela J., ed. *Policing the Black Man: Arrest, Prosecution, and Imprisonment*. Pantheon Books, 2017.

Davis, Angela Y. *Are Prisons Obsolete?* Seven Stories Press, 2003.

del Pozo, Brandon. *The Police and the State: Security, Social Cooperation, and the Public Good*. Princeton University Press, 2022.

Deshpande, Manasi, and Michael Mueller-Smith. "Does Welfare Prevent Crime? The Criminal Justice Outcomes of Youth Removed from SSI." *Quarterly Journal of Economics* 137, no. 4 (2022): 2263–307.

Du Mont, Janice, Karen-Lee Miller, and Terri L. Myhr. "The Role of 'Real Rape' and 'Real Victim' Stereotypes in the Police Reporting Practices of Sexually Assaulted Women." *Violence Against Women* 9, no. 4 (2003): 466–86.

Eck, John, and R. T. Guerette. "Place-Based Crime Prevention: Theory, Evidence, and Policy." In *The Oxford Handbook of Crime Prevention*, edited by Brandon C. Welsh and David P. Farrington. Oxford University Press, 2012.

Foley, C. Fritz. "Welfare Payments and Crime." *Review of Economics and Statistics* 93, no. 1 (2011): 97–112.

Freedman, Matthew, and Emily G. Owens. "Low-Income Housing Development and Crime." *Journal of Urban Economics* 70, no. 2 (2011): 115–31.

Friedman, Barry. "Disaggregating the Police Function." *University of Pennsylvania Law Review* 169, no. 4 (2021): 925–1002.

Fryer, Daniel. "Idealizing Abolition." *Criminal Law and Philosophy* 17, no. 3 (2023): 553–72.

Ghandnoosh, Nazgol, and Celeste Barry. "One in Five: Disparities in Crime and Policing." Sentencing Project, November 2023. https://www.sentencingproject.org/reports/one-in-five-disparities-in-crime-and-policing/.

Goldsmith, Andrew. "Police Reform and the Problem of Trust." *Theoretical Criminology* 9, no. 4 (2005): 443–70.

Gray, Sharon. "Community Safety Workers: An Exploratory Study of Some Emerging Crime Prevention Occupations." International Centre for the Prevention of Crime, 2006.

Guerrero, Alexander. "Against Elections: The Lottocratic Alternative." *Philosophy and Public Affairs* 42, no. 2 (2014): 135–78.

Hannon, Lance, and James DeFronzo. "Welfare and Property Crime." *Justice Quarterly* 15, no. 2 (1998): 273–88.

Heath, Joseph. "The Challenge of Policing Minorities in a Liberal Society." *Journal of Political Philosophy* (2023).

Herskind, Micah, and Tiffany Roberts. "The Failure of Police Reform: Atlanta's City Government Spent Decades on Half-Hearted Measures Meant to Address Police Violence." *Intellingencer* (blog). *New York Magazine*, January 31, 2022. https://nymag.com/intelligencer/2022/01/atlanta-police-reform-failure.html.

Hunt, Luke William. "The Limits of Reallocative and Algorithmic Policing." *Criminal Justice Ethics* 41, no. 2 (2022): 21–44.

———. *Police Deception and Dishonesty: The Logic of Lying.* Oxford University Press, 2023.

———. "Policing, Brutality, and the Demands of Justice." *Criminal Justice Ethics* 40, no. 1 (2021): 40–55.

Jones, Ben. "Police-Generated Killings: The Gap Between Ethics and the Law." *Journal of Political Philosophy* 29, no. 3 (2021): 345–67.

Kaba, Mariame. "Yes, We Mean Literally Abolish the Police." *New York Times*, June 12, 2020.

Kaba, Mariame, and Andrea J. Ritchie. *No More Police: A Case for Abolition*. New Press, 2022.

Karma, Roge. "4 Ideas to Replace Traditional Police Officers." *Vox*, June 24, 2020.

Kelley, Robin D. G. "What Abolition Looks Like, from the Panthers to the People." *Level* (blog). Medium, October 26, 2020. https://level.medium.com/what-abolition-looks-like-from-the-panthers-to-the-people-6c2e537eac71.

Kirk, David S., and Sara Wakefield. "Collateral Consequences of Punishment:

A Critical Review and Path Forward." *Annual Review of Criminology* 1, no. 1 (2018): 171–94.
Kleinig, John. *The Ethics of Policing*. Cambridge University Press, 1996.
Lai, Calvin K., and Jaclyn A. Lisnek. "The Impact of Implicit-Bias-Oriented Diversity Training on Police Officers' Beliefs, Motivations, and Actions." *Psychological Science* 34, no. 4 (2023): 424–34.
Landemore, Hélène. *Democratic Reason*. Princeton: Princeton University Press, 2012.
Levin, Sam. "US Police Have Killed Nearly 600 People in Traffic Stops Since 2017, Data Shows." *Guardian*, April 21, 2022.
Macaré, Niels, Maya Schenwar, and Alana Yu-lan Price, eds. *Who Do You Serve, Who Do You Protect? Police Violence and Resistance in the United States*. Haymarket Books, 2016.
Maher, Geo. *A World Without Police: How Strong Communities Make Cops Obsolete*. Verso Books, 2021.
Mapping Police Violence. "2024 Police Violence Report." 2024. https://policeviolencereport.org/.
McCorkel, Jill. "The Second Coming: Gender, Race, and the Privatization of Carceral Drug Treatment." *Contemporary Drug Problems* 44, no. 4 (2017): 286–300.
Miller, Eric J. "The Concept of the Police." *Criminal Law and Philosophy* 17, no. 3 (2023): 573–95.
Mills, Charles W. "Ideal Theory as Ideology." *Hypatia* 20, no. 3 (2005): 165–84.
Modestino, Alicia Sasser. "How Do Summer Youth Employment Programs Improve Criminal Justice Outcomes, and for Whom?" *Journal of Policy Analysis and Management* 38, no. 3 (2019): 600–28.
Monaghan, Jake. "Broken Windows, Naloxone, and Experiments in Policing." *Social Theory and Practice* 48, no. 2 (2022): 309–30.
———. "Idealizations and Ideal Policing." *Philosophers' Imprint* 22, no. 9 (2022).
———. *Just Policing*. Oxford University Press, 2023.
———. "Legitimate Policing and Professional Norms." In *The Ethics of Policing: New Perspectives on Law Enforcement*, edited by Ben Jones and Eduardo Mendieta. New York University Press, 2021.
———. "The Special Moral Obligations of Law Enforcement." *Journal of Political Philosophy* 25, no. 2 (2017): 218–37.
Mowen, Thomas, and John Brent. "School Discipline as a Turning Point: The Cumulative Effect of Suspension on Arrest." *Journal of Research in Crime and Delinquency* 53, no. 5 (2016): 628–53.
Nagin, Daniel S., Robert M. Solow, and Cynthia Lum. "Deterrence, Criminal Opportunities, and Police." *Criminology* 53, no. 1 (2015): 74–100.

National Center for Health Statistics. "Firearm Mortality by State, 2022." Centers for Disease Control and Prevention, 2022. https://www.cdc.gov/nchs/pressroom/sosmap/firearm_mortality/firearm.htm.

Page, Jennifer. "Defensive Killing by Police: Analyzing Uncertain Threat Scenarios." *Journal of Ethics and Social Philosophy* 24, no. 3 (2023): 315–51.

———. "Reparations for Police Killings." *Perspectives on Politics* 17, no. 4 (2019): 958–72.

Powers, Madison, and Ruth Faden. *Structural Injustice: Power, Advantage, and Human Rights*. Oxford University Press, 2019.

Purnell, Derecka. *Becoming Abolitionists: Police, Protests, and the Pursuit of Freedom*. Astra House, 2021.

Ralph, Laurence. *The Torture Letters: Reckoning with Police Violence.* University of Chicago Press, 2019.

Ratcliffe, Jerry H. "Policing and Public Health Calls for Service in Philadelphia." *Policing* 43, no. 4 (2020): 643–58.

Rawls, John. *A Theory of Justice*. Harvard University Press, 1971.

Ristroph, Alice. "Just Violence." *Arizona Law Review* 56, no. 4 (2014): 1017–46.

Ritchie, Andrea J. *Invisible No More: Police Violence Against Black Women and Women of Color*. Beacon Press, 2017.

Sharkey, Patrick. "Community and the Crime Decline: The Causal Effect of Local Nonprofits on Violent Crime." *American Sociological Review* 83, no. 6 (2018): 1214–40.

———. "Why Do We Need the Police?" *Washington Post*, June 12, 2020.

Sierra-Arévalo, Michael. *The Danger Imperative: Violence, Death, and the Soul of Policing*. Columbia University Press, 2024.

Sorensen, Lucy C., Yinzhi Shen, and Shawn D. Bushway. "Making Schools Safer and/or Escalating Disciplinary Response: A Study of Police Officers in North Carolina Schools." *Educational Evaluation and Policy Analysis* 43, no. 3 (2021): 495–519.

Spencer, Patricia C., and Shari Munch. "Client Violence Toward Social Workers: The Role of Management in Community Mental Health Programs." *Social Work* 48, no. 4 (2003): 532–44.

Stroebe, Wolfgang, Pontus Leander, and Arie W. Kruglanski. "Gun Ownership and Gun Violence: A Comparison of the United States and Switzerland." *Aggression and Violent Behavior* 78, no. 1 (2024): 101987.

Subramanian, Ram, and Leily Arzy. "State Policing Reforms Since George Floyd's Murder." Brennan Center for Justice, May 21, 2021. https://www.brennancenter.org/our-work/research-reports/state-policing-reforms-george-floyds-murder.

Thacher, David. "Shrinking the Police Footprint." *Criminal Justice Ethics* 41, no. 1

(2022): 62–85.
Thompson, Derek. "Unbundle the Police." *Atlantic*, June 11, 2020.
Thompson, Jamie. "To Stop a Shooter." *Atlantic*, January 29, 2024.
Fuller, Doris A., H. Richard Lamb, Michael Biasotti, and John Snook. *Overlooked in the Undercounted: The Role of Mental Illness in Fatal Law Enforcement Encounters*. Treatment Advocacy Center, 2015. https://www.tac.org/reports_publications/overlooked-in-the-undercounted-the-role-of-mental-illness-in-fatal-law-enforcement-encounters/.
Venema, Rachel M. "Police Officer Schema of Sexual Assault Reports: Real Rape, Ambiguous Cases, and False Reports." *Journal of Interpersonal Violence* 31, no. 5 (2016): 872–94.
Vera Institute of Justice, "Behavioral Health Crisis Alternatives: Shifting from Police to Community Responses." November 2020. https://www.vera.org/behavioral-health-crisis-alternatives.
Vitale, Alex S. *City of Disorder: How the Quality of Life Campaign Transformed New York Politics*. New York University Press, 2008.
———. *The End of Policing*. Verso Books, 2021.
Vogler, Jacob. "Access to Health Care and Criminal Behavior: Short-Run Evidence from the ACA Medicaid Expansions." November 2017. http://dx.doi.org/10.2139/ssrn.3042267.
Wacquant, Loïc. *Punishing the Poor: The Neoliberal Government of Social Insecurity*. Duke University Press, 2009.
Webster, John A. "Police Task and Time Study." *Journal of Criminal Law, Criminology, and Political Science* 61, no. 1 (1970): 94–100.
Weisburst, Emily K. "Patrolling Public Schools: The Impact of Funding for School Police on Student Discipline and Long-Term Education Outcomes." *Journal of Policy Analysis and Management* 38, no. 2 (2019): 338–65.
Western, Bruce. "The Impact of Incarceration on Wage Mobility and Inequality." *American Sociological Review* 67, no. 4 (2002): 526–46.
World Population Review. "Gun Ownership by State." World Population Review, 2025. https://worldpopulationreview.com/state-rankings/gun-ownership-by-state.
Wuschke, Kathryn, Martin A. Andresen, and Paul Brantingham. "What Do Police Do and How Do They Do It? A Systematic Review of Police Time-Use Studies." *International Journal of Police Science and Management* 20, no. 1 (2018): 19–27.
Young, Iris Marion. *Responsibility for Justice*. Oxford University Press, 2013.
Zacka, Bernardo. *When the State Meets the Street: Public Service and Moral Agency*. Harvard University Press, 2017.
Zimring, Franklin E. *When Police Kill*. Harvard University Press, 2018.

Journal of Ethics and Social Philosophy
VOL. 30, NO. 3 · MAY 2025
https://doi.org/10.26556/jesp.v30i3.4324

IN DEFENSE OF THE TROLLEY METHOD

Ian Stoner and Jason Swartwood

GUY CRAIN recently argued that the trolley method of moral philosophy has three shortcomings that have not yet been appreciated by its practitioners.[1] To the extent that Crain's criticisms highlight ways fanciful examples are sometimes abused, we welcome them. We are moved to reply because Crain suggests that he has identified shortcomings of fanciful examples as they are routinely employed by philosophers. We disagree.

1. WHAT FANCIFUL EXAMPLES ARE FOR

Our fundamental objection to Crain's approach is that his characterization of fanciful examples in moral philosophy neglects one of their key features: fanciful examples, competently employed as philosophical tools of persuasion or critical reflection, always have the goal of shaping the beliefs of readers in some specific way.[2]

Fanciful examples play a variety of belief-shaping roles. They ground arguments from analogy, function as paradigm cases for abducting moral principles, and present scenarios that invite conceptual clarification. But perhaps the most common use of fanciful examples (and of described cases in general) occurs within the method of wide reflective equilibrium, where they "function as either *counter-examples* or *reductio ad absurda*."[3] Authors develop examples intended to elicit from readers a reflectively endorsed moral judgment about a concrete case that is clearly in tension with a specific target belief. Used this way, fanciful examples are a "custom-built tool for illuminating the ill-fittingness of a target belief."[4]

1 Crain, "Three Shortcomings of the Trolley Method of Moral Philosophy," (hereafter cited parenthetically).

2 We prefer the umbrella term 'fanciful examples' for any wholly invented, heavily stipulated described case. The metonymic 'trolley case' is needlessly misleading; Crain's and others' criticisms of this method cover many examples that do not feature trolleys.

3 Walsh, "A Moderate Defence of the Use of Thought Experiments in Applied Ethics," 471.

4 Stoner and Swartwood, "Fanciful Examples," 326.

In this discussion note, we focus on fanciful examples used as counterexamples. The well-known cases Crain catalogs in the introduction to his article are all arguably deployed as counterexamples to specific target beliefs. Consider three of the fanciful examples he highlights.

Organ Transplant targets the belief that it is morally permissible to kill one person in order to save five people.[5] Judith Jarvis Thomson's audience is readers who are tempted by the original trolley problem to abstract the belief that the lives of the many outweigh the life of one. Thomson expects that such readers will judge that it is *not* morally permissible for a surgeon to kill one healthy person in order to transplant his organs into five desperate patients, and this judgment is obviously in tension with the belief that it is permissible to kill a person in order to save many others.

Ticking Time Bomb targets the belief that torture is always, in principle, wrong.[6] Those who have used this example expect that readers will form the judgment that torture is permissible in situations in which it is the only way to save many innocent lives, and this belief is in tension with the belief that torture is categorically impermissible.

Jim the Botanist targets the belief that ethical thinking is a simple algorithm that takes as input all and only those levers of power currently available to the agent.[7] Bernard Williams expects readers will find it hard to form a judgment about what Jim should do when Pedro offers to spare the lives of nineteen condemned innocents if Jim personally executes one of the twenty captives. Even though there is only one lever of power available to Jim, and the consequences of pulling it are clear, the decision Jim faces is not trivially easy to settle.

Attempts to deploy fanciful examples as counterexamples fail when they fail to elicit a reflectively endorsed judgment about the described case that is in tension with the target belief. There are many available paths to such failure. Some examples dictate stipulations that are difficult or even impossible for most readers to imagine.[8] Some examples invite unreliable judgments that arise from framing effects or bigoted background beliefs.[9] Failures such as these are mistakes in applying the method, not weaknesses of the method itself.

Crain's target is the method itself. Some of Crain's objections are prudential objections, others are methodological. Our position is that Crain's objections

5 Thomson, "Killing, Letting Die, and the Trolley Problem," 206. Thomson's Footbridge example (207–8) is structured the same way and targets the same belief.

6 Alhoff, "Ticking Time-Bombs and Torture."

7 Smart and Williams, *Utilitarianism*, 98–99.

8 McGee and Foster, *Intuitively Rational*, 161–63.

9 Wood, "Humanity as an End in Itself," 69; and Stoner and Swartwood, "Fanciful Examples," 330–33.

all fail, for the underlying reason that he has neglected the belief-shaping goal of fanciful examples. Once we acknowledge that examples aim to shape beliefs in specific ways, it is clear that the objections Crain develops are not objections to the use of fanciful examples but objections to their abuse.

2. Methodological Objection: Trolley Cases Fail to Model Real-World Ethical Decisions

In the real world, most ethical decisions are influenced by past experiences and feature limited conscious reflection; real-world decisions usually occur under some degree of uncertainty and rarely involve discrete decision points with enumerated options. Real ethical decisions, in other words, are nothing like trolley problems. Crain suggests that this failure to model the conditions of real-world decisions undermines the trolley method as a method of moral philosophy (433–34).[10]

The trouble with this objection is that it misrepresents the methodological role of fanciful examples. They are not intended to model ordinary moral decision-making. They are intended to test specific moral beliefs that an audience holds using that audience's own judgments about described cases. The goal of examples is to improve the set of moral beliefs that we will, in turn, employ in real-world decision-making.

Consider an analogy to the difference between medical research and everyday medical practice. Imagine a randomized controlled trial designed to test a new cancer drug. That trial will yield new information about the effectiveness of the drug, and that information will in turn shape the treatment plans oncologists develop for their patients. The researchers' study of drug efficacy is, in an important way, involved in practical decision-making about treatment, but only when clinicians apply in the practical domain the information generated by the research study. Different methods are appropriate for determining the general efficacy of the drug and deciding whether the drug is appropriate for a particular patient in a particular situation, and it would be absurd for a clinician to object that randomized controlled trials are methodologically flawed on the grounds that they fail to model the complicated, fluid decisions clinicians make in partnership with their patients in choosing treatment plans.

Similarly, philosophers who use fanciful examples aim to improve the set of beliefs we employ in practice. Just as with medical research and medical practice, these are two different areas of inquiry. There is no reason to expect that

10 See also Wood, "Humanity as an End in Itself," 70; and Fried, "What Does Matter?" 506.

philosophical methods that refine a set of beliefs should mirror the application of that set of beliefs in practical decision-making.

Take, for example, Peter Unger's trolley case, Bob's Bugatti.[11] Unger's target is the belief that charity's demands must be limited—people who make significant sacrifices on behalf of strangers may deserve praise, but their sacrifices are supererogatory. Against this belief, Unger sketches a situation in which a man invests his retirement savings in a Bugatti. He parks his retirement Bugatti just beyond a railroad siding and goes for a hike. On returning, he spots a child trapped on the main track in the path of an oncoming train. Bob can direct the train onto the siding, which will destroy his retirement Bugatti, or he can do nothing, in which case his retirement investment will remain secure, but the child will be crushed. Bob chooses to let the child die. In response to this case, nearly everyone forms the judgment that Bob's decision is monstrous.[12]

Bob's Bugatti does not model typical real-world decisions; Crain is right that we rarely or never find ourselves facing choices like Bob's. But Unger has targeted a belief that is involved in familiar practical decisions about charitable giving. Someone who believes that morality cannot demand much in the way of charity will think differently about how to set up their monthly budget than someone who does not hold that belief. Just as a randomized controlled trial can help a clinician improve their beliefs about the efficacy of the treatments they prescribe, Bob's Bugatti can help us improve our beliefs about the limits of our charitable obligations.

When careless thinkers attempt to treat fanciful examples as models for real-world choices, disaster sometimes ensues. This has been hashed out at length in discussions of Ticking Time Bomb. Ticking Time Bomb, properly employed, can demonstrate that hardly anyone believes, on reflection, that torture is always, in principle, wrong. But it is a gross abuse of that case to argue that because torture is the right thing to do in a wildly unrealistic fantasia, torture should be an option available to real-world governments.[13] People who have defended real-world torture by appeal to Ticking Time Bomb are reasoning poorly *because* they are treating the fanciful example as a model of real-world decision-making when, like most fanciful examples, it is not and should not be.

11 Unger, *Living High and Letting Die*, 136.

12 In addition to functioning as a counterexample, Bob's Bugatti can play other belief-shaping roles, such as functioning as a paradigm case for an inference to the best explanation (which is one of the roles it plays in Unger's original) or as the basis of an argument from analogy (which is how Peter Singer uses the same case in "The Singer Solution to World Poverty").

13 Beck and de Wijze, "Interrogating the 'Ticking Bomb Scenario.'"

3. Methodological Objection: Because They Involve High Stakes, Trolley Cases Are Useless for the Ethics of the Mundane

Fanciful examples are typically extreme and often violent; they involve death and dismemberment, earthquakes and boat accidents, puppy murders and disintegrator rays. The questions most of us face in daily life are nothing like that. Should I call my mom? Should I microwave my fish sticks in the employee break room? May I pass along a juicy morsel of gossip, or must I leave it unplucked? This gulf between the enormity of fanciful examples and the mundanity of the ethical choices most of us actually face leads Crain to object that "the trolley method is, by design, a terrible tool for working on the ethics of the mundane" (427).

If it were true that the extreme features of fanciful examples make them useless for mundane ethics, then that would be a mark against the method. But in many familiar domains—perhaps most famously, physics and mathematics—extreme cases, carefully employed, can help focus and correct the beliefs and patterns of reasoning that we apply to realistic cases.

Take, for example, a classic puzzle of mathematics, the Monty Hall problem. Monty Hall presents you with three doors, only one of which hides a valuable prize. You make your selection—say, door three. Monty Hall then opens one of the other doors—say, door one—revealing that there is nothing behind it. He then gives you the opportunity to change your guess. Should you stick with door three or change your guess to door two?

When first presented with the Monty Hall problem, many people's statistical judgments are mistaken—they believe that the odds are even that they have picked the winning door, and switching their guess would be pointless. One way to make clear that this belief is false is to imagine an extreme version of the case. Suppose Monty Hall presents you with one hundred doors, and behind one of them is a prize. You guess door three. He then opens ninety-eight of the remaining doors, revealing nothing behind them—only door three and door two remain closed. Should you stick with door three or switch your guess to door two? In this case, it is obvious that you should switch your guess. Recognizing this makes it easy to see that in the real-world, three-door version of the game, you should likewise switch your guess. The extreme case helps us achieve a clarity of understanding that helps us navigate the initially murkier real-world case.

Fanciful examples in moral philosophy can play a similar role. When it comes to our obligations of charity to strangers, some of our moral beliefs are arguably mistaken. Bob's Bugatti is an extreme case intended to help. It shines a light on a target belief that many of us have adopted from our families and communities: that significant charitable giving is supererogatory. By providing

us with a fanciful example that defamiliarizes charitable giving, Unger makes it easy for us to see that Bob's choice to prioritize his own secure retirement is not, by our own lights, morally permissible; significant levels of financial sacrifice *can* be obligatory. Recognizing this makes it easier to see that the mundane question of how to structure our monthly budget is morally fraught.

Bob's Bugatti is not an outlier. Other fanciful examples are structurally similar. They often target beliefs that are centrally implicated in the ethics of the mundane. They often accomplish their goal by providing extreme cases that allow us to see those beliefs in a clearer, because less familiar, light.[14] Fanciful examples are obviously not the only tool for working on everyday moral problems, but they can help improve beliefs relevant to the ethics of the mundane. The fact that extreme cases in any domain—physics, mathematics, and ethics, too—can be misleading when used carelessly is a reason to use them carefully, not to reject them categorically.

4. Methodological Objection: Faux Anonymization of Agents Invites Distorted Responses

Crain notes that in most fanciful examples, "but for the sparsest of features, the agent is anonymized" either via minimally sketched third-person cases or second-person cases in which each reader is explicitly invited to imagine themselves as the agent (430). Crain suggests that anonymization is problematic because readers supply their own details for these minimally described cases, and regardless of their backgrounds, readers tend to supply details that track closely with Western, educated, industrialized, rich, and democratic (WEIRD) cultural assumptions (429). The result of faux anonymization is that non-WEIRD agents "are likely being operationally excluded.... The trolley method is not ethics for everyone" (431).

Crain's suggestion that non-WEIRD readers supply WEIRD details for anonymized agents runs counter to our own classroom experiences. But suppose he is right. Suppose readers, regardless of their own backgrounds, tend to imagine WEIRD agents in anonymized cases—most everyone, and not just our mothers, picture George Clooney when invited to imagine Thomson's transplant surgeon. This would not undermine the value of Organ Transplant as a counterexample to the target belief that it is morally permissible to kill one to save many. Anyone who imagines George Clooney in Organ Transplant still

14 Walsh, "A Moderate Defence of the Use of Thought Experiments in Applied Ethics," 473–74.

has before them a concrete case in which they themselves judge that it would be wrong to kill one to save many.

The same is true for other fanciful examples employed in their various belief-shaping roles. Examples can be effective in their persuasive roles even if non-WEIRD readers imaginatively supply WEIRD cultural details, so long as those readers respond to the example in the way the author intends. If WEIRD beliefs prompt reactions to the case that are morally misleading or incorrect, that merely provides additional beliefs the method can target for revision, not a reason to reject the method.

In our classrooms, we have noticed that students, including students from non-WEIRD communities, often supply details drawn from their own experiences when presented with under-described cases written by other students. It sometimes happens that the culturally specific details they supply flip the valence of the judgment the case's author intended to elicit. That is a marker of a case that is *problematically under-described*.[15] An effective counterexample must be broadly accessible and uncontroversial.[16] If anonymization of agents typically results in examples that are so under-described that they are unable to accomplish their persuasive role, then that would constitute a methodological criticism of anonymization. But there is no reason to think that anonymization of agents typically risks problematic under-description of a case. Note that in every one of Crain's opening examples, including Organ Transplant, Jim the Botanist, and Ticking Time Bomb, agents are anonymized, and yet these are famous examples in part *because* most readers, regardless of their backgrounds, form in response to them the judgments that their authors intend to elicit.[17]

In most cases, anonymization of agents is a kindness authors extend to their readers. Anonymization makes it easier for us all to imagine *ourselves* in the situation the author has sketched, with all of our existing values, beliefs, backgrounds, and experiences. Consider a de-anonymized version of Thomson's Organ Transplant:

> *Sara the Transplant Surgeon*: Sara is a thirty-seven-year-old divorced mother of two. She was raised in a conservative Muslim home, and

15 This is the case in Crain's own Bystander case (429). That deliberately under-described case lacks sufficient detail for readers to judge whether intervention is warranted.

16 Stoner and Swartwood, *Doing Practical Ethics*, ch. 3.

17 To be clear, eliciting a common judgment from a diverse audience is not sufficient for these or any other examples to succeed as counterexamples. The elicited judgment must withstand reflective scrutiny and be in genuine tension with the target belief. It might turn out, after further argument, that everyone's judgment about the case is misguided and should be revised. We simply intend to highlight that anonymization of agents does not undermine the effectiveness of otherwise well-designed cases.

> though she still observes some practices, she has quietly come to think of herself as agnostic. Her relationship with her parents was strained by her divorce, but over the previous few months they have made efforts to extend olive branches to each other. Sara finds her work as a surgeon rewarding, but she is buffeted by regrets: of her huge outstanding student loans, of her curtailed time with her children, of her status as perpetually on call. She worries that she bonds with her patients more than other surgeons do and that she may be substituting connection to her patients for her attenuated connections to her parents and children. Should Sara murder one innocent bystander in order to save five of her patients?

Sara the Transplant Surgeon reads as a gentle parody because it is clear that her biography is irrelevant to Thomson's project. What matters to Thomson is that *you*, reader, with *your* beliefs, *your* values, *your* biography, judge that it would be wrong to kill an innocent in order to harvest their organs.[18] Part of what makes Thomson's example effective is that readers' judgments about it are not contingent on their biographical idiosyncrasies. WEIRD and non-WEIRD alike, most readers agree that harvesting organs is no excuse for murder.

5. PRUDENTIAL OBJECTIONS

In addition to his methodological objections, Crain worries about the impression fanciful examples leave on students and the public. Fanciful examples, with their characteristically high stakes "build the impression that ethics just amounts to the rare decision faced by many or the normal decision faced by the few" (427). Relatedly, since it is lower-stakes decisions that "make up the stuff of everyday moral life" (426), fanciful examples contribute to the impression that philosophers are uninterested in the ethical issues most people face. When authors anonymize agents, that opens the possibility that they will "insidiously create the impression that to be a moral agent at all is to be a WEIRD, young, and fit male" (432). These prudential objections persist even if we are right that Crain's methodological objections fail. But we wish to highlight that prudential objections provide no reason to avoid the method of fanciful examples in philosophical research.

18 We thank an anonymous reviewer for pointing out that the best characterization of Thomson's method remains an area of interpretative debate. See, e.g., Conte, "The Trolley Problem and Intuitional Evidence." Our argument does not rest on whether Thomson would agree that we have correctly characterized her method. We expect she would agree that her anonymization of the surgeon in Organ Transplant does not undermine the effectiveness of the example.

Philosophy teachers and public philosophers should recognize that philosophical methods, including the method of fanciful examples, are unfamiliar to many people. Audiences unfamiliar with the trolley method—whether students or members of the public—should be supplied with the background they need to understand its means and goals. Poor communication of the trolley method, in classrooms and in public venues, risks doing harm.[19]

But just as the analogous challenges of science education and science communication do not give scientists reason to abandon scientific methods, the challenges of philosophical education and public philosophy do not give us reason to abandon philosophical methods. We should do our best to teach our students and to educate the public about the role fanciful examples play in refining the set of moral beliefs that shape our everyday moral life. If the method is useful, as we have argued it is, then the risk of public misunderstanding gives us reason to work harder at teaching well, not reason to abandon the method.

6. Conclusion

We have argued that Crain's objections to the trolley method of moral philosophy can be defused by taking seriously the belief-shaping role of fanciful examples. We have focused on fanciful examples as counterexamples to specific target beliefs, but our arguments can easily be adapted for fanciful examples used in other specific belief-shaping roles: as analogies, paradigm cases, and so on. Competently employed in pursuit of the goal of shaping beliefs in these ways, fanciful examples can be extraordinarily useful philosophical tools.

Although Crain's objections fail as criticisms of standard uses of the trolley method, his discussion serves to highlight an important way examples can fail. When authors, teachers, and public figures fail to identify the specific beliefs their fanciful examples are designed to shape, they flirt with several follies.[20]

Saint Paul College
ian.stoner@saintpaul.edu
jason.swartwood@saintpaul.edu

19 Martena, "Thinking Inside the Box."

20 We are grateful to three anonymous reviewers for comments that improved this discussion note.

REFERENCES

Allhoff, Fritz. "Ticking Time-Bombs and Torture." In *Contemporary Debates in Applied Ethics*, edited by Andrew I. Cohen and Christopher Heath Wellman. 2nd ed. John Wiley and Sons, 2013.

Beck, Simon, and Stephen de Wijze. "Interrogating the 'Ticking Bomb Scenario': Reassessing the Thought Experiment." *International Journal of Applied Philosophy* 29, no. 1 (2015): 53–70.

Conte, Sebastian J. "The Trolley Problem and Intuitional Evidence." *Journal of Value Inquiry* 58, no. 2 (2024): 293–310.

Crain, Guy. "Three Shortcomings of the Trolley Method of Moral Philosophy." *Journal of Ethics and Social Philosophy* 26, no. 2 (2023): 420–43.

Fried, Barbara. "What Does Matter? The Case for Killing the Trolley Problem (or Letting It Die)." *Philosophical Quarterly* 62, no. 248 (2012): 505–29.

Martena, Laura. "Thinking Inside the Box: Concerns About Trolley Problems in the Ethics Classroom." *Teaching Philosophy* 41, no. 4 (2018): 381–406.

McGee, Andrew, and Charles Foster. *Intuitively Rational: How We Think and How We Should*. Springer Nature, 2024.

Singer, Peter. "The Singer Solution to World Poverty." *New York Times Magazine*, September 5, 1999.

Smart, J. J. C., and Bernard Williams. *Utilitarianism: For and Against*. Cambridge University Press, 1973.

Stoner, Ian, and Jason Swartwood. *Doing Practical Ethics*. Oxford University Press, 2021.

———. "Fanciful Examples." *Metaphilosophy* 48, no. 3 (2017): 325–44.

Thomson, Judith Jarvis. "Killing, Letting Die, and the Trolley Problem," *Monist* 59, no. 2 (1976): 204–17.

Unger, Peter. *Living High and Letting Die*. Oxford University Press, 1996.

Walsh, Adrian. "A Moderate Defence of the Use of Thought Experiments in Applied Ethics." *Ethical Theory and Moral Practice* 14, no. 4 (2011): 467–81.

Wood, Allen. "Humanity as an End in Itself." In Derek Parfit, *On What Matters,* vol. 2. Oxford University Press, 2011.

JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY
http://www.jesp.org
ISSN 1559-3061

The *Journal of Ethics and Social Philosophy* (JESP) is a peer-reviewed online journal in moral, social, political, and legal philosophy. The journal is founded on the principle of publisher-funded open access. There are no publication fees for authors, and public access to articles is free of charge. Articles are typically published under the CREATIVE COMMONS ATTRIBUTION-NONCOMMERCIAL-NODERIVATIVES 4.0 license, though authors can request a different Creative Commons license if one is required for funding purposes.

Funding for the journal has been made possible through the generous commitment of the Division of Arts and Humanities at New York University Abu Dhabi.

جامعة نيويورك أبوظبي
NYU ABU DHABI