

The Ethics of Civil Commitment

By Jonathan Cantarero*

Involuntary confinement is the most serious deprivation of individual liberty that a society may impose. The philosophical justifications for such a deprivation by means of the criminal process have been thoroughly explored. No such intellectual effort has been directed at providing justifications for societal use of civil commitment procedures.¹

Caring for individuals with mental illness was historically understood as a household issue.² When families failed to act, states often jailed those they deemed to be “lunatics” in the name of public safety.³ The 19th and 20th centuries, however, saw an increase use of asylums and psychiatric centers, which, coupled with the lax commitment laws of the time, made it easy to commit individuals involuntarily.⁴ While some viewed this as a sign of growing public concern for the mentally ill,⁵ others saw it as proof of more cynical motives foreshadowing the eugenics movement.⁶

* J.D., Assistant District Attorney, Manhattan District Attorney’s Office. The author would like to thank Professor Sofia Yakren, of CUNY School of Law, for her assistance in preparing this article. All views expressed herein are solely those of the author.

¹ Joseph M. Livermore, et al., *On the Justifications for Civil Commitment*, 117 U. PA. L. REV. 75 (1968). While this quote precedes most of the disability rights movement, the literature on philosophical critiques of mental health law continues to lag in comparison to other legal areas such as criminal law.

² See FRED PELKA, WHAT WE HAVE DONE: AN ORAL HISTORY OF THE DISABILITY RIGHTS MOVEMENT 48 (2012) (noting most mentally ill individuals historically were not institutionalized); JUDITH L. FALLER, WHO QUALIFIES FOR RIGHTS? 72 (2002) (noting families tended to keep disabled members at home unless extreme circumstances existed).

³ See PELKA *supra* note 2, at 48. In 1788, New York provided statutory authority for precisely these types of cases:

Whereas, there are sometimes persons who by lunacy or otherwise are furiously made, or are so far disordered in their senses that they be dangerous to be permitted to go abroad . . . it shall and may be lawful for any two or more justices of the peace to cause such person to be apprehended and kept safely locked up in some secure place, and if . . . necessary, to be there chained.

Id. (citing *New York Laws of 1788*, chap. 31, quoted in Deutsch, *Mentally Ill in America*, 419).

⁴ See Lauren Hoopes, *On the Periphery: A Survey of Nineteenth-Century Asylums in the United States* (May 2015) (unpublished graduate thesis, All Theses, Clemson University) (noting increased number of asylums and negative effects of Civil War contributed to involuntary commitment).

⁵ See PELKA, *supra* note 2, at 49 (noting majority opinion that isolation and control in institutions were best for “feeble minded”).

⁶ See *id.* at 49 (noting formation of disability rights activists’ groups).

In either case, the mass commitment that ensued led the majority of states to significantly heighten standards for commitment.⁷ Accordingly, at present, most states now require demonstration—in addition to a showing of mental illness—that an individual is dangerous to themselves or others before such commitment may be warranted.⁸ Regardless of the state, however, the law surrounding civil commitment has always been based on two distinct but equally important sources of state power: the state acting as *parens patriae* (“parent of the country”) in caring for the individual and the state acting through its police power in protecting society.⁹

This article discusses the ethical justification for the current standard for involuntary civil commitment of individuals with mental illness from a philosophical ethical perspective. To begin, I suggest that determining whether the state is acting as *parens patriae* or through its police powers, is wholly dependent on how the state chooses to apply the standard.¹⁰ For example, the state acts as *parens patriae* insofar as it seeks to confine a person who is “mentally ill” and “dangerous to *self*,” but acts under its police powers when the individual is alleged to be “mentally ill” and “dangerous to *others*.”¹¹ Second, I argue that while these state powers authorize involuntary civil commitment under specific criteria, each form of state action is in turn justified under corresponding ethical principles: *Parens patriae* by Kantian ethics and police power through utilitarianism.¹² Finally, I contend that the balancing of these ethical theories as reflected in the current standard provides us with a morally defensible standard for civil commitment.¹³ This, I argue, is because the common flaws traditionally associated with each ethical principle, are largely offset when they work side-by-side rather than in isolation.¹⁴

I. Introduction

In the landmark involuntary commitment case *Addington v. Texas*, the Supreme Court discussed the government’s authority to commit individuals for mental health treatment:

The state has a legitimate interest under its *parens patriae* powers in providing care to its citizens who are unable because of emotional disorders to care for themselves; the state also has authority under its

⁷ See Megan Testa & Sara G. West, *Civil Commitment in the United States*, 7 PSYCHIATRY (EDGMONT) 30, 32 (2010) (noting abuses of civil commitment resulted in shift in standards).

⁸ See generally, BRUCE W. WINICK, CIVIL COMMITMENT 75-81 (2005) (providing state-by-state definitions for mental illness or mental disability for civil commitment). Other states have used comparable terms such as “gravely disabled” in lieu of “mentally ill.” *Id.*

⁹ See MICHAEL L. PERLIN, THE HIDDEN PREJUDICE: MENTAL DISABILITY ON TRIAL 79 (2000) (revealing pattern of prejudice against mentally disabled individuals which prevents equal treatment).

¹⁰ See *infra* Part III.

¹¹ See, e.g., Developments in the Law, *Civil Commitment of the Mentally Ill*, 87 HARV. L. REV. 1190, 1207-09 (1974) (discussing origins and nature of *parens patriae*).

¹² See *infra* Part IV.

¹³ See *infra* Part V.

¹⁴ See *infra* Part V.

police power to protect the community from the dangerous tendencies of some who are mentally ill.¹⁵

In order to lay the framework for an ethical assessment of involuntary civil commitment, there are three facets of the problem that need to be distinguished: The power, the purpose, and the ethical justification. *Parens patriae* and police power doctrines are commonly understood as justifications for civil commitment.¹⁶ To avoid confusion, however, this article uses those principles as merely sources of authority. The above quote from the *Addington* Court, for example, declares that the state has an interest under its *parens patriae* power and has *authority* under its police power. The point in both cases is that the state has the *ability* to act, not that it is necessarily *correct* in so doing.¹⁷ To make the latter moral judgment, we first need to look at the specific purpose for exercising each power. In other words, each purpose must be based on some adequate ethical principle in order to be considered morally defensible. A prerequisite, however, is to define the terms at issue.

II. Defining the Terms

On the heels of the disability rights movement, the 1970's saw a remarkable heightening of protection for those with mental illness.¹⁸ Following the Supreme Court's decision in *O'Connor*¹⁹ and the federal District Court for the Eastern District of Wisconsin's reasoning in *Lessard*,²⁰ many state statutes began requiring—in addition to a showing of a “mental illness”—“dangerousness to self or others,” before commitment was lawful.²¹ Though these terms have become amorphous and difficult to define, I

¹⁵ *Addington v. Texas*, 441 U.S. 418, 426 (1979). The court further concluded that an individual's interest in the outcome of a civil commitment proceeding is of such weight and gravity that due process requires the state to justify confinement by proof more substantial than a mere preponderance of the evidence. *Id.* at 427.

¹⁶ See, e.g., Seth Feuerstein et al., *Civil Commitment, A Power Granted to Physicians by Society*, 2 PSYCHIATRY (EDGMONT) 53, 54 (2005) (discussing *parens patriae* and police power).

¹⁷ On this point, the Supreme Court has previously held that “the fact that state law may have authorized confinement of the harmless mentally ill does not itself establish a constitutionally adequate purpose for the confinement.” *O'Connor v. Donaldson*, 422 U.S. 563, 574 (1975) (emphasis added); Geoffrey E. Linburn, *Donaldson Revisited: Is Dangerousness a Constitutional Requirement for Civil Commitment?* 26 J. AM. ACAD. PSYCHIATRY L. 343, 346 (1998) (arguing that common interpretation that dangerousness is required actually misreads case holding).

¹⁸ See W. Lawrence Fitch et al., *Civil Commitment and the Mental Health Care Continuum: Historical Trends and Principles for Law and Practice*, SUBSTANCE ABUSE AND MENTAL HEALTH SVCS ADMIN. 37 (2019), https://www.samhsa.gov/sites/default/files/civil-commitment-continuum-of-care_041919_508.pdf (discussing origins of mental health system and shift towards deinstitutionalization).

¹⁹ See *O'Connor*, 422 U.S. at 575-76. “A finding of ‘mental illness’ alone cannot justify a State's locking a person up against his will ... the State cannot constitutionally confine without more a nondangerous individual who is capable of surviving safely in freedom by himself or with...help.”

²⁰ See *Schmidt v. Lessard*, 349 F. Supp. 1078, 1095 (E.D. Wis. 1972), *vacated*, 414 U.S. 473 (1974) (“We therefore hold that the state must prove beyond a reasonable doubt all facts necessary to show that an individual is mentally ill and dangerous.”). The Supreme Court vacated and remanded the case on procedural grounds. *Id.* at 477.

²¹ See WINICK, *supra* note 8, at 75-97.

mitigate this issue by extracting only those general principles which are pertinent to this analysis.

Beginning with “mental illness,” courts tend to assume rather than assess this requirement in civil commitment statutes. Indeed, a brief survey of the relevant case law shows us that a disproportionate amount of text is devoted, not to assessing mental illness, but to examining dangerousness.²² Commentators note that the threshold finding of mental illness required in such cases is all too often viewed as being “sufficiently obvious as to not be a matter of discussion.”²³ As a consequence, courts often defer to both the legislature’s statutory construction and the clinician’s testimony of mental illness to support an affirmative finding in that regard.²⁴ Notably, while deference to the legislature in applying a statute is readily expected, deference to the medical field is more likely attributable to the fact that the term mental illness implicates medical expertise in a way that dangerousness does not. Arrigo identifies this judicial-clinical syncretism as a sort of clinicolegal jurisprudence (i.e., cases where medical expertise effectively substitutes legal judgment).²⁵

The problem with judicial deference to state statutes, however, is that state statutes have historically been replete with broad and circular definitions of mental illness.²⁶ Texas, for example, previously defined “mental illness” tautologically as “a person whose mental health is substantially impaired.”²⁷ More recent statutes are similarly vague and broad.²⁸ Alabama, for example, currently defines mental illness as “a psychiatric disorder of thought and/or mood which significantly impairs judgment, behavior, [and the] capacity to recognize reality or ability to cope with the ordinary demands of everyday life.”²⁹

²² See generally *Donahue v. R.I. Dep’t of Mental Health, Retardation & Hosp.*, 632 F. Supp. 1456, 1478 (D. R.I. 1986) (discussing how courts applied *O’Connor’s* dangerousness requirement); *Hubbart v. Superior Court*, 969 P.2d 584, 599-601 (Cal. 1999) (upholding statute permitting confinement based on potential for repeat sexually violent offenses under dangerousness analysis); *In re Doe*, 78 P.3d 341, 356-62 (Haw. Ct. App. 2003) (discussing federal dangerousness cases and struggles to define standard because Court established no bright-line rule); see also, *supra* notes 15, 17, and accompanying text.

²³ PERLIN, *supra* note 9, at 79.

²⁴ See *United States v. Comstock*, 627 F.3d 513, 518 (4th Cir. 2010) (discussing judicial deference to congressional judgment when federal statute is attacked); *United States v. Brandon*, 158 F.3d 947, 952 (6th Cir. 1998) (discussing judicial deference to professional medical judgment); *Bailey v. Gardebring*, 940 F.2d 1150, 1153 (8th Cir. 1991) (discussing federal court deference to Minnesota court’s interpretation of state civil commitment statute).

²⁵ BRUCE A. ARRIGO, PUNISHING THE MENTALLY ILL: A CRITICAL ANALYSIS OF LAW AND PSYCHIATRY 83-86 (2012).

²⁶ *Id.* at 75-81.

²⁷ See, e.g., TEX. REV. CIV. STAT. ANN., art. 5547-4(k) (West 1982-83) (repealed 1991); John Kip Cornwell, *Understanding the Role of the Police and Parens Patriae Powers in Involuntary Civil Commitment Before and After Hendricks*, 4 PSYCHOL. PUB. POL’Y & L. 377, 387, n.93 (1998) (citing Texas and Colorado statutory definitions).

²⁸ See ARRIGO, *supra* note 25, at 84-85.

²⁹ ALA. CODE § 22-52-1.1(1) (2014). Similar examples appear in the current statutes of Alaska, Colorado, and Washington. Alaska defines mental illness as “an organic, mental, or emotional impairment that has substantial adverse effects on an individual’s ability to exercise conscious control of the individual’s actions or ability to perceive reality or to reason or understand.” ALASKA STAT. § 47.30.915 (2002). Alaska’s statute further states that “intellectual disability,

Consistent with an understanding of mental illness as a legal term, Alabama³⁰ and other states³¹ consciously narrowed the scope of the definition to distinguish disorders of thought and mood from other impairments related to developmental disabilities and substance abuse.³² This method of construction echoes Winick's view that the question of mental illness for the purpose of civil commitment is a purely legal question and that—given the great deprivation of liberty at stake—it should be defined in a narrow sense.³³

It suffices for our purposes to deduce three basic principles to guide our understanding of mental illness. First, to meet the mental illness requirement, the individual must have some cognizable psychological impairment. Second, this impairment must have a current negative effect on the individual. Third, the effect must be severe enough to substantially jeopardize his or her welfare.

In contrast to mental illness, dangerousness as a requirement for involuntary commitment was only re-introduced in the 1970s in order to remedy the mass commitments which occurred in large part because of the lax standard that preceded it.³⁴ Vague by its nature, this term is often referred to as one of the few elusive doctrines in mental disability law.³⁵ We can nevertheless break down dangerousness into three components: (1) The type of action covered; (2) how proximate in time the action may be expected to occur; and (3) the evidence required to show that such action will in fact take place. Contrary to the minimal case law dedicated to defining mental illness, courts have scrutinized dangerousness in greater depth, shaping the definition in several landmark opinions.³⁶

developmental disability, or both, epilepsy, drug addiction, and alcoholism do not per se constitute mental illness, although persons suffering from these conditions may also be suffering from mental illness.” *Id.* Colorado’s definition is “one or more substantial disorders of the cognitive, volitional, or emotional processes that grossly impairs judgment or capacity to recognize reality or to control behavior.” COLO. REV. STAT. § 27-65-102 (2003). In Washington, “mental disorder” means “any organic, mental, or emotional impairment which has substantial adverse effects on a person’s cognitive or volitional functions.” WASH. REV. CODE ANN. § 71.05.020(36) (LexisNexis 2003).

³⁰ *Id.*

³¹ *See, e.g., In re Michael S.*, 636 N.Y.S.2d 261, 263 (Sup. Ct. 1995) (declining to equate mental illness with drug addiction in an involuntary civil commitment proceeding).

³² *See, e.g.,* MINN. STAT. § 253B.09 (2014) (authorizing commitment based on mental illness, developmental disability and substance abuse); MINN. STAT. §§ 253B.02(2), (13)-(14) (2014) (defining and providing different criteria for chemical dependence, mental illness and developmental disabilities). *See also* Justin C. Wilson, *Case Comment: Will Full Benefits Parity Create Real Parity?: Congress’s Second Attempt at Ending Discrimination Against Mental Illness: The Paul Wellstone and Pete Domenici Mental Health and Addiction Equity Act of 2008*, 3 ST. LOUIS U. J. HEALTH L. & POL’Y 343, 373 (discussing state parity laws passed before and after the Mental Health Parity Act).

³³ *See* WINICK, *supra* note 8, at 48. “What should constitute mental illness for civil commitment purposes is a legal, not a clinical question.” *Id.* “Because civil commitment results in a deprivation of the fundamental liberty interest in being free of external restraint, a narrow definition of mental illness is required.” *Id.*

³⁴ *See* ARRIGO, *supra* note 25, at 87; Fitch et al., *supra* note 18, at 3-9.

³⁵ *See* ARRIGO, *supra* note 25, at 86; PERLIN, *supra* note 9, at 80.

³⁶ *See* cases cited *supra* notes 20, 22, 24.

Beginning with the types of actions covered, the New Jersey Supreme Court in *Krol* loosely equated dangerousness with a risk of “harm” to others.³⁷ The court noted that this harm, while often understood as contemplating a *physical* injury,³⁸ can also take the form of *psychological* trauma³⁹ or even destruction of property.⁴⁰ Second, with regard to proximity, the court in *Lessard* described the requirement as “an extreme likelihood that if the person is not confined he will do *immediate* harm to himself or others.”⁴¹ In doing so, the court equated the term with *imminent* danger.⁴² Third, *Lessard* noted that, while not necessary, “a recent overt act, attempt, or threat to do substantial harm to oneself or another,” could provide sufficient evidence to establish dangerousness.⁴³ As a result, some courts now not only favor, but require an “overt act” as a prerequisite for dangerousness—although many states have declined to push the standard that far.⁴⁴ Given this variety, it is unsurprising that state statutes have largely mimicked the discord in case law.⁴⁵ One aspect that is clear, however, is the standard of proof required. The Supreme Court in *Addington v. Texas* expressly rejected a “beyond reasonable doubt” standard in favor of the more generous “clear and convincing evidence” standard for proving dangerousness.⁴⁶ This is of course consistent with the basic understanding that involuntary commitment is a civil process rather than a criminal procedure.⁴⁷

A few basic principles may be gleaned from these cases. Most obviously, the case law demonstrates that dangerousness is defined in terms of a concern for the well-being both of the subject individual and others. Second, whether raised as a concern to self or others, the danger at issue is generally equated with *harm*. This harm, in turn, can be physical (to a person or property), or psychological. Third, the closer the proximity and the greater the gravity of the harm, the more likely it is that a court will find a person to be dangerous for purposes of civil commitment. Here, past acts only serve as indicators of future conduct and are not, in and of themselves, grounds for

³⁷ *State v. Krol*, 344 A.2d 289, 301 (N.J. Sup. Ct. 1975).

³⁸ *See, e.g., In re Harry M.*, 96 A.D.2d 201, 206 (N.Y. App. Div. 1983) (limiting the definition of harm as a “substantial threat to...*physical* well-being”) (emphasis added).

³⁹ *See Krol*, 344 A.2d at 301.

⁴⁰ *See Id.*

⁴¹ *Lessard v. Schmidt*, 349 F. Supp. 1078, 1093 (D. Wis. 1972); *see also Suzuki v. Yuen*, 617 F.2d 173, 178 (9th Cir. 1980) (citing *Lessard*, 349 F. Supp. at 1093 and using the term *imminent*).

⁴² *See Lessard*, 349 F. Supp. at 1094. “The necessity of a finding of imminent danger to oneself or others is strengthened by a comparison to persons incapacitated by physical illness.” *Id.*

⁴³ *Id.* at 1093.

⁴⁴ *See Stamus v. Leonhardt*, 414 F. Supp. 439, 451 (S.D. Iowa 1976) (requiring “overt act, threat or attempt”); *but see U. S. ex rel. Mathew v. Nelson*, 461 F. Supp. 707, 708 (N.D. Ill. 1978) (finding that absence of overt act does not violate due process); *Fisk v. Letterman*, 501 F. Supp. 2d 505, 524 (S.D.N.Y. 2007) (declining to require overt act).

⁴⁵ *See WINICK*, *supra* note 8, at 81-86.

⁴⁶ *See Addington v. Texas*, 441 U.S. 418, 431-33 (1979). “We noted earlier that the trial court employed the standard of ‘clear, unequivocal and convincing’ evidence in appellant’s commitment hearing before a jury. That instruction was constitutionally adequate.” *Id.* at 433.

⁴⁷ *See Donald Stone, There Are Cracks in the Civil Commitment Process: A Practitioner’s Recommendations to Patch the System*, 43 *FORDHAM URB. L.J.* 789, 790-791 (2016) (discussing civil commitment process when person is deemed dangerously mentally ill). “A criminal defendant is often guaranteed greater protections than a mentally ill person facing involuntary civil commitment.” *Id.*

commitment.⁴⁸ With these working definitions in mind, we are in a better position to understand the sources of power that bring these standards to life.

III. *Parens Patriae* and Police Power

The Supreme Court states two chief but distinct sources of authority through which a state may involuntarily commit an individual: *Parens patriae*⁴⁹ and police power.⁵⁰ The state acting as *parens patriae* generally refers to the state's inherent power and authority to act in the best interest of those individuals who lack the capacity to do so.⁵¹ In the context of mental health, this translates to the government's power to care for individuals who, as a result of mental illness, are unable to care for themselves.⁵² This doctrine confers power not only to *protect* but to *control* the individual in need of care.⁵³

Courts have upheld the state's *parens patriae* powers in a wide variety of settings.⁵⁴ Because of the great loss of personal freedom, however, this power is not unlimited. The Supreme Court in *O'Connor* underscored that the use of *parens patriae* authority requires a balancing of the state's interest and the deprivation at stake.⁵⁵ In the context of involuntary commitment, the state's interest in protecting one member of its community is weighted against the individual's interest in personal autonomy.⁵⁶ In *O'Connor*, the Supreme Court ultimately found that the petitioner had a constitutionally

⁴⁸ See *Stamus*, 414 F. Supp. at 450-451 (outlining standards for civil commitment). "Due process and equal protection require that the standards for commitment must be (a) that the person is mentally ill and poses a serious threat of substantial harm to himself or to others; and (b) that this threat of harm has been evidenced by a recent overt act or threat." *Id.* See also Karen Franklin, *The Best Predictor of Future Behavior Is ... Past Behavior*, PSYCHOLOGY TODAY (Jan. 3, 2013), <https://www.psychologytoday.com/us/blog/witness/201301/the-best-predictor-future-behavior-is-past-behavior> (discussing how past behavior is an indicator for future conduct only under certain circumstances).

⁴⁹ See, e.g., *Addington*, 441 U.S. at 426 (1979). Here, the Court outlines the state's interests and abilities under its *parens patriae* and police powers, within the context of addressing citizens with emotional disorders. *Id.* With respect to its *parens patriae* powers, the Court points to the state's responsibility as a caretaker of those who cannot care for themselves. *Id.*

⁵⁰ See *id.* For police powers, the Court highlights the state's responsibility in ensuring that its communities are safe from those with more dangerous tendencies, due to mental illness. *Id.* See also FALLER, *supra* note 2, at 68.

⁵¹ See Developments in the Law, *supra* note 11, at 1208. Historically, *parens patriae* referred to the role of the King – under English law during the American colonial period – as the *de facto* guardian of certain special classes of people; "all persons who had lost their intellects and become ... incompetent to take care of themselves." *Id.* (quoting *In re Barker*, 2 Johns. Ch. 232, 236 (N.Y. Ch. 1816)). For a discussion on the origins of the doctrine of *parens patriae*, see generally Lawrence B. Custer, *The Origins of the Doctrine of Parens Patriae*, 27 EMORY L. J. 195 (1978).

⁵² See FALLER, *supra* note 2, at 68.

⁵³ See John Seymour, *Parens Patriae and Wardship Powers: Their Nature and Origins*, 14 OXFORD J. LEGAL STUD. 159, 160 (1994).

⁵⁴ See, e.g., *Late Corp. of Church of Jesus Christ v. United States*, 136 U.S. 1, 65-66 (1890) (stating property may revert to state under *parens patriae*); *In re Eugene W.*, 29 Cal. App. 3d 623 (Cal. Ct. App. 1972) (affirming state's interest to defend welfare of minors under *parens patriae*); *O'Connor v. Donaldson*, 422 U.S. at 583 (reasserting states' rights to civilly commit people with mental illnesses).

⁵⁵ See *O'Connor*, 422 U.S. at 580-81; see also Developments in the Law, *supra* note 11, at 1208-11.

⁵⁶ See *O'Connor*, 422 U.S. at 582-83.

protected right to liberty.⁵⁷ As a result, it held that “the State [could not] constitutionally confine without more a[n] individual who is capable of surviving safely in freedom by himself or with the help of willing and responsible family members or friends.”⁵⁸ In *Addington*, the Supreme Court reiterated this rationale, noting that justification under *parens patriae* grounds requires “showing that the individual suffers from something more [than] idiosyncratic behavior.”⁵⁹ Similarly, *parens patriae* power does not justify commitment if the only purpose for confinement is treatment.⁶⁰

Looking at its use more closely, *parens patriae*, when applied in commitment cases, authorizes two key aspects of the involuntary commitment standard. First, it functions as a means for assisting those who are so mentally ill as to endanger their own welfare.⁶¹ This corresponds to the mental illness requirement of the standard.⁶² Second, it validates the infringement on personal liberty by framing the application as one which protects those individuals from harming themselves, likewise linking itself to the danger to self requirement.⁶³ This application thus functions to discharge the state from its duty to its most vulnerable.⁶⁴ Because the power stems from its interest in protecting the person rather than society as a whole, *parens patriae* is fundamentally an individual-oriented source of authority.

On the other end of the spectrum is another inherent power of government. It has been said that no term has been so successful in escaping definition as “police power.”⁶⁵ To be sure, courts routinely support the state’s police power but seldom articulate what it actually means or encompasses.⁶⁶ One scholar writing during the era of mass commitment attempted to describe it as “an exercise of the unclassified, residuary power of government, vested in the state to promote the public welfare of its inhabitants.”⁶⁷ I think, however, that police power is most easily viewed in contradistinction to our *parens patriae* definition. From this perspective, police power means simply the state’s inherent power to protect *society* from potential harm.⁶⁸

⁵⁷ See *id.* at 576.

⁵⁸ *Id.* at 576.

⁵⁹ *Addington v. Texas*, 441 U.S., 418, 427 (1979).

⁶⁰ See *O’Connor*, 422 U.S. at 583-84 (Burger, C.J., concurring).

⁶¹ See *Addington*, 441 U.S. at 426; see also *Youngberg v. Romeo*, 457 U.S. 307, 324-25 (1982) (holding states must provide services to committed mental patients as necessary to ensure their safety).

⁶² See *Addington*, 441 U.S. at 426. I do not assume that all persons with mental illness cannot care for themselves. The narrow definitions in state statutes likewise reflect this view by ensuring additional criteria in order for a person to meet the “legal definition.” See *supra* text accompanying notes 13-17.

⁶³ See *Addington*, 441 U.S. at 426; see also *O’Connor*, 422 U.S. at 576 (holding state cannot confine non-dangerous individual capable of safely living on own without more).

⁶⁴ See FALLER, *supra* note 2, at 70; see also *Gregory v. Rodgers*, 974 F.2d 1006, 1010 (8th Cir. 1992) (noting Due Process imposes duty to protect citizens with limited ability to care for themselves).

⁶⁵ See Walter W. Cook, *What is the Police Power?* 7 COLUM. L. REV. 322, 322-336 (1907) (discussing several formulations and definitions for police power from early twentieth century).

⁶⁶ *Id.* at 322.

⁶⁷ *Id.* at 329.

⁶⁸ ROBERT M. LEVY & LEONARD S. RUBENSTEIN, *THE RIGHTS OF PEOPLE WITH MENTAL DISABILITIES: THE AUTHORITATIVE ACLU GUIDE* 17 (1st ed., 1996).

Police power has a long, if not longer history than that of *parens patriae* and was similarly invoked in a variety of contexts.⁶⁹ Indeed, many of the earlier common law statutes for imprisoning the incompetent were premised on the state's police power.⁷⁰ Recent cases have likewise echoed the principle that the state's police power authorizes it to commit an individual if doing so would protect society.⁷¹ It thus has deep roots in our legal system.

Just as *parens patriae* corresponds to a specific version of the current standard for involuntary commitment, police power goes hand-in-hand with another. That is to say where a state wishes to confine an individual based on his or her dangerousness to others, the state's police powers are readily implicated. The police power in the simplest sense seeks to protect a large group of people residing within the state. Likewise, the interest in confining a person with a mental illness due to the risks he or she poses to others allows the state to use its inherent power to protect others from individuals with a cognizable mental illness. The key distinction, therefore, is that where the state has acted as *parens patriae*, it has generally used its power in the purported interest of a particular person, namely, the individual being committed. Similarly, where the state utilized its police power, it has fundamentally done so in the purported common interest of society. The ultimate question, however, is whether either use of inherent state power is ethical.

IV. Ethics of Kant and Mill

That the state acts under particular inherent sources of authority (either as *parens patriae* or under police power), depending on whether such action is in the interest of a person or society, is uncontroversial. Less clear, however, are the moral assumptions behind such actions in the case of involuntary civil commitment. While analysis of these moral assumptions is underdeveloped in the relevant literature, these assumptions provide much needed support for the current standard for involuntary commitment. Moreover, because involuntary commitment is fundamentally about a person's unwilling loss of liberty, it is critical to evaluate its moral underpinnings. This section lays out the foundational background for utilitarianism and Kantian ethics. It then suggests how these ethical theories justify state police power and *parens patriae* power, specifically within the context of involuntary commitment. Finally, it addresses critiques of each theory and rebuttals to those arguments.

⁶⁹ See Cook *supra* note 65, at 326-28; Santiago Legarre, *The Historical Background of the Police Power*, 9 U. PA. J. CONST. L. 745, 781-96 (2007) (discussing origins of police power doctrine).

⁷⁰ See *supra* notes 2-3 and accompanying text.

⁷¹ See, e.g., *In re Necessity for the Hospitalization of Naomi B.*, 435 P.3d 918, 931 (Alaska 2019) ([W]hen a person has been found likely to cause harm to others . . . the State has a compelling interest in protecting the public, grounded in its police power.); *People v. Torski C.* (In re Torski C.), 918 N.E.2d 1218, 1230 (Ill. App. Ct. 2009) ("Although the State's goal in protecting society from harm certainly justifies police-power action, whether the State's interest is great enough to support imposition of the deprivations associated with civil commitment depends on the nature of the threat posed by the mentally ill individual."); *In re A.J.N.*, 144 A.3d 130, 137 (Pa. Super. Ct. 2016) ("The government's authority to involuntarily commit persons who are a danger to themselves or others arises from the state's inherent police powers.").

A. Utilitarianism and Police Power

Utilitarianism is a “normative” ethical theory, dealing not with what *is* the case but what *ought* to be the case.⁷² As such, utilitarianism seeks to establish a standard by which to guide and appraise human action.⁷³ Its founder, Jeremy Bentham, stated in his “Principle of Utility” that “it is the greatest happiness of the greatest number that is the measure of right and wrong.”⁷⁴ In *An Introduction to the Principles of Morals and Legislation*, Bentham added:

Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. They alone point out what we *ought* to do and determine what we *shall* do; the standard of right and wrong, and the chain of causes and effects, are both fastened to their throne...[t]he principle of utility recognises this subjection and makes it the basis of a system that aims to have the edifice of happiness built by the hands of reason and of law.⁷⁵

Bentham’s student, John Stuart Mill, adopted Bentham’s maxim and incorporated the above reasoning in his own version of the Principle of Utility:

The creed which accepts as the foundations of morals “utility” or the “greatest happiness principle” holds that actions are right in proportion as they tend to promote happiness; wrong as they tend to produce the reverse of happiness. By happiness is intended pleasure and the absence of pain; by unhappiness, pain and the privation of pleasure.⁷⁶

Both versions agree that the test for whether an action is right or wrong is whether it produces happiness or unhappiness. Similarly, both define happiness in relation to pain and pleasure. If we agree with Bentham and Mill on how to distinguish right from wrong, and we assume that we should do right rather than wrong, it follows that we should act only to maximize happiness for the greatest number.

In addition to this, however, later scholars have advocated not only for the maximization of happiness, but the minimization of unhappiness through the newer doctrine of “Negative Utilitarianism.”⁷⁷ As formulated by British philosopher Karl

⁷² See DAVID LYONS, *ETHICS AND THE RULE OF LAW* 111 (3d ed. 1984).

⁷³ See *id.* at 111, 117, 127.

⁷⁴ JEREMY BENTHAM, *A FRAGMENT ON GOVERNMENT* 93 (F. C. Montague ed., Oxford: Clarendon Press 1881, reprinted by The Lawbook Exchange, Ltd. 2001) (1776) (stating Bentham’s fundamental axiom).

⁷⁵ Jeremy Bentham, *The Principle of Utility*, in *AN INTRODUCTION TO THE PRINCIPLES OF MORALS AND LEGISLATION* 6 (Jonathan Bennett trans., 2017) (1780), available at <http://www.earlymoderntexts.com/assets/pdfs/bentham1780.pdf> (emphasis in original).

⁷⁶ John Stuart Mill, *Utilitarianism*, in *CLASSICS OF WESTERN PHILOSOPHY* 1063, 1066-67 (Steven M. Cahn ed., Hackett Publishing Co. 3d ed. 1990) (1863).

⁷⁷ See Roderick Ninian Smart, *Negative Utilitarianism*, 67 *MIND* 542, 542-43 (1958) (using phrase “negative utilitarianism” to discuss philosopher Karl Popper’s formulation); see also Harry B. Acton & John W. N. Watkins, *Negative Utilitarianism*, 37 *ARISTOTELIAN SOC’Y SUPPLEMENTAL VOL.* 83, 83-114 (1963).

Popper, and later termed “Negative Utilitarianism” by R. N. Smart, the concept is understood “as the responsibility to *prevent* the greatest amount of harm or evil.”⁷⁸ Popper used Negative Utilitarianism to extend Mill’s logic and require affirmative steps to minimize unhappiness. Similarly, Richard Hull incorporated Popper’s logic into what he called the “Principle of Beneficence” or the “duty to produce good, prevent harm, and remove harm.”⁷⁹

To be sure, Mill alluded to these subsequent developments in utilitarian philosophy in his classic work *On Liberty*.⁸⁰ There, Mill sought to address the ethical limits of restrictions by the government on personal autonomy.⁸¹ In doing so, he extended the Principle of Utility by incorporating the “Harm Principle,” which advocated inaction to prevent unhappiness.⁸²

[T]he only purpose for which power can be rightfully exercised over any member of a [civilized] community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear, because it will be better for him to do so, because it will make him happier, because, in the opinions of others, to do so would be wise or even right. These are good reasons for remonstrating with him, or reasoning with him, or persuading him, or entreating him, but not for compelling him or visiting him with any evil in case he do otherwise.⁸³

This view was based on the notion that when the state uses its power over an individual against his or her will, the resulting action is a restriction on the person’s autonomy – consent, protest or desire becomes immaterial.⁸⁴ Moreover, when that power prevents the individual from taking certain action, it becomes a restriction on individual liberty.⁸⁵

⁷⁸ See Karl R. Popper, 1 *THE OPEN SOCIETY AND ITS ENEMIES* 65, n. 6 (Princeton Univ. Press 4th ed., 1963). Popper did not use the phrase “negative utilitarianism” per se, but instead explained:

I suggest, for this reason to replace the utilitarian formula ‘Aim at the greatest happiness for the greatest number’ . . . by the formula ‘The least amount of avoidable suffering for all,’ or briefly, ‘Minimize suffering.’ Such a simple formula can, I believe, be made one of the fundamental principles . . . of public policy.

Id. See also Smart, *supra* note 77, at 542-43.

⁷⁹ See, e.g., Richard T. Hull, *Involuntary Commitment and Treatment of Persons Diagnosed as Mentally Ill*, *BIOMEDICAL ETHICS REV.* 131, 135-40 (James M. Humber & Robert F. Almedar eds., Humana Press 1983), http://www.richard-t-hull.com/publications/involuntary_commitment.pdf.

Articulating the risk of harm posed to society at large is a major justification for the commitment of mentally ill individuals. *Id.* at 140. The principle is used to illustrate the idea that society feels that commitment of such individuals will not only protect the individual themselves and have a positive impact on their lives, but will also have a positive impact on society at large. *Id.*

⁸⁰ See John Stuart Mill, *On Liberty*, in *CLASSICS OF WESTERN PHILOSOPHY* 1105, 1164-78 (Steven M. Cahn ed., Hackett Publishing Co. 3d ed. 1990) (1869).

⁸¹ See *id.* at 1110-11.

⁸² See *id.*

⁸³ *Id.*

⁸⁴ See *id.*

⁸⁵ *Id.* at 1110-12.

While the prior maxims did not explicitly weigh individual liberty against utilitarian considerations, Mill's Harm Principle clearly allows the latter to outweigh the former. Given Mill's well-known interest in protecting individual liberty,⁸⁶ his acknowledgment that such liberty may nevertheless be restricted in certain circumstances is telling. In formulating the Harm Principle, he did not argue, for example, that liberty could *never* be restricted; but rather, he recognized a higher interest in protecting the greatest number of people from harm, even when that meant inflicting harm on an individual.⁸⁷

Relating utilitarianism to involuntary civil commitment, it is important to note that the connection between traditional utilitarianism and the state's police power has been previously acknowledged.⁸⁸ Mill claimed that a restrictive action is justified only if it prevents harm to *others* (plural)⁸⁹—not another (singular). That position is consistent with the state's use of its police power to act in the interest of society at large. In addition, scholars interpret the Harm Principle as referring to what types of conduct the *state* is justified in restricting,⁹⁰ thereby setting up an ethical standard under which the state's police power may operate.⁹¹

Negative Utilitarianism further imposes on the state a duty to engage in affirmative action minimizing unhappiness to the greatest number—rather than engaging in mere inaction under the Principle of Utility.⁹² Thus, the state's use of its police power, to minimize unhappiness caused by one individual, is generally justified under the Negative Utilitarian framework outlined above.

In addition, to Negative Utilitarianism's justification of the state's police power, scholars also point specifically to the connection between utilitarianism and involuntary civil commitment.⁹³ Hull, for example, views Mill's Harm Principle as “the basic argument” for the second version of involuntary commitment— where the individual is

⁸⁶ See Mill, *supra* note 80, at 1110-12; see also See e.g. Mark Strasser, *Mill and the Utility of Liberty*, 34 PHIL. Q. 63 (1984) (“Yet it is precisely because Mill was a champion of liberty as a utilitarian that he deserves our praise and respect.”) (emphasis omitted).

⁸⁷ David O. Brink, *Mill on Justice and Rights*, A COMPANION TO MILL 374, (Christopher Macleod & Dale E. Miller eds., Wiley-Blackwell 2016), <https://davidobrink.com/sites/default/files/publications/MillJusticeRightsProof.pdf> (last visited Mar. 9, 2020).

⁸⁸ See generally *Limiting the State's Police Power: Judicial Reaction to John Stuart Mill Source*, 37 U. CHI. L. REV. 605 (1970) (discussing Mill's influence on “police power” jurisprudence in United States).

⁸⁹ See Mill, *supra* note 80, at 1111. While it is conceivable that a person civilly confined is a danger to only one other individual (e.g. a schizophrenic bent on assassinating the president), Mill and subsequent scholars uniformly understood “others” to mean society in general and thus refers to multiple people (e.g., if some person is bent on assassinating the president, he or she is likely willing to harm others in the process). *Id.*

⁹⁰ See *Limiting the State's Police Power: Judicial Reaction to John Stuart Mill*, *supra* note 88, at 608 (discussing role of harm principal in judicial review of laws).

⁹¹ See generally, *id.*

⁹² See Smart, *supra* note 77, at 542-43 (defining negative utilitarianism); *supra* note 76 and accompanying text (stating Mill's Principle of Utility).

⁹³ See generally Livermore, *supra* note 1, and Hull, *supra* note 79, at 137-39.

“mentally ill” and “a danger to others.”⁹⁴ Hull, however, adds the Principle of Beneficence⁹⁵ to justify affirmative steps taken by the state in depriving someone of liberty via involuntary commitment.⁹⁶ In other words, the state has a duty under its police powers to minimize unhappiness. This duty prevails even if it results in the deprivation of liberty to a single individual, so long as the net harm and pain prevented (the guideposts for happiness and unhappiness) is greater than the net harm and pain produced by involuntary commitment.⁹⁷ When prevention of pain outweighs its production, involuntary civil commitment is an ethical means of minimizing unhappiness to the greatest number of people.⁹⁸

Livermore et al., support this reasoning on limited terms. While they argue that utilitarianism does not justify the state in confining a person who is *mentally ill* and merely *unduly burdensome* to others,⁹⁹ they nevertheless acknowledge that confining someone who is *dangerous to others* serves a higher interest than mere “burdens” or “nuisances,” and can justify involuntary commitment.¹⁰⁰ Because Mill himself conceded that harm to others, and not mere nuisances, was a sufficient basis for a deprivation of one individual’s liberty,¹⁰¹ and given that neither version of the standard for involuntary civil commitment allows for confinement on the basis of mere nuisances, we can assume the interest of preventing danger to others is great enough to meet Mill’s requirements. Indeed the very definition of danger to others in the case law is often articulated as risk of *harm* to others.¹⁰²

Though it appears that we have, in Negative Utilitarianism, an ethical theory to defend the state’s practice of involuntary commitment where the individual is *mentally ill*

⁹⁴ Hull, *supra* note 79, at 139 (discussing why criminal punishment inadequate to prevent harm to others so civil commitment is beneficial).

⁹⁵ See *id.* at 140 (discussing the principle of beneficence). See also TOM BEAUCHAMP AND JAMES CHILDRESS, PRINCIPLES OF BIOMEDICAL ETHICS (1st ed. 1980).

⁹⁶ See Hull, *supra* note 79, at 140.

⁹⁷ See generally Hull, *supra* note 79. The “harm principle” is when the harm is caused or would be caused. *Id.* at 140. “It is morally justified to prevent harm to [other] persons when the harm is caused or would be caused by those whose liberty is restricted.” *Id.* at 135 (quoting BEAUCHAMP & CHILDRESS, *supra* note 95). This principle rests on the “principle of beneficence, understood as a duty to produce good, prevent harm, and remove harm.” *Id.* at 140.

⁹⁸ See *id.* at 141. Hull wrote:

[S]ince the harm principle justifies restriction of liberty only when doing so prevents harm, if the only negative consequences that can be foreseen accrue just to the agent, and the agent has accepted their possibility, those consequences would not be harms whose possibility could serve as the basis for invoking the harm principle.

Id.

⁹⁹ See Livermore, *supra* note 1, at 87 (illustrating justification as a “utility of confinement” versus “utility of liberty” issue).

¹⁰⁰ See *id.* at 88. “Our libertarian views usually lead us to assert that treatment cannot be forced on anyone unless the alternative is very great *social harm*” (emphasis added). *Id.*

¹⁰¹ See generally Mill, *supra* note 80.

¹⁰² See, e.g., State v. Krol, 344 A.2d 289, 301 (N.J. Sup. Ct. 1975). “The risk of danger, a product of the likelihood of such conduct and the degree of harm which may ensure, must be substantial within the reasonably foreseeable future.” *Id.* at 302.

and *dangerous to others*, Utilitarianism is often criticized for being, in principle, a consequence-based theory or one that gives no weight to the motive of actors.¹⁰³ In other words, under traditional utilitarianism, actions are right as long as they tend to *result* in greater aggregate happiness, irrespective of whether the actor's motives are hedonistic or even malicious. Under this theory, it appears irrelevant whether the state's motive in committing individuals against their will is either to eliminate them from the general population or to protect a township from imminent danger—Utilitarians would say each is equally valid. Indeed, Bentham states that, “[i]f the [motives] are good or bad, it is only on account of their effects[.]”¹⁰⁴ This, however, begs the question: Why do “motives” matter?

To begin, “the motive of an action is the ultimate desire of the agent that explains its occurrence, or some feature of it.”¹⁰⁵ As Steven Sverdlik has previously suggested, motives matter because they have the potential for intrinsic value.¹⁰⁶ He argued, for example, that refusing to shake someone's hand to avoid spreading a cold was very different than refusing to do so because they are of a different race.¹⁰⁷

Foucault's “Social Control Thesis” develops this critique and places it within the context of civil commitment.¹⁰⁸ Foucault argued that “confinement of the noncriminal is a method of controlling the socially undesirable.”¹⁰⁹ He reasoned that the use of institutions for confinement of people with mental illness was a means of policing “public hygiene” and served ultimately to rid society of the socially undesirable.¹¹⁰ Similarly, interpretations of Foucault's theory view involuntary civil commitment as essentially “state-sanctioned social control,”¹¹¹ echoing the same concerns raised by the initial rise of asylums—that a quasi-eugenic movement meant to control and contain outcasts in society was underway.¹¹² The chief complaint with respect to involuntary civil commitment then is that a traditionally utilitarian State is committed to such action so long as its effect is greater overall happiness (or, in our

¹⁰³ See Mill, *supra* note 80. Recall that Mill's principle of utility states that “actions are right in proportion as they tend to *promote* happiness; wrong as they tend to *produce* the reverse of happiness.” Mill, *supra* note 76, at 1066-67 (emphasis added); David Brink, *Mill's Moral and Political Philosophy*, THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Winter 2018 Ed.), <https://plato.stanford.edu/entries/mill-moral-political/> (last visited Mar. 1, 2020).

¹⁰⁴ See Robert Merrihew Adams, *Motive Utilitarianism*, 73 J. PHIL. 467, 478 (1976) (citing Jeremy Bentham, *in Happiness Is the Greatest Good*).

¹⁰⁵ STEVEN SVERDLIK, MOTIVE AND RIGHTNESS 18 (2011).

¹⁰⁶ *Id.* at 55.

¹⁰⁷ See generally, *id.* at chapter 5.

¹⁰⁸ See ARRIGO, *supra* note 25, at 83-84 (discussing Foucault's social control thesis). See also MICHEL FOUCAULT, DISCIPLINE AND PUNISH, THE BIRTH OF PRISON (Alan Sheridan trans., Vintage Books 2d ed. 1995) (1977) (explaining society has the right to punish under this view); MICHEL FOUCAULT, MADNESS AND CIVILIZATION: A HISTORY OF INSANITY IN THE AGE OF REASON (Richard Howard trans., Pantheon Books 1964) (1961).

¹⁰⁹ See ARRIGO, *supra* note 25, 76 (contending civil commitment and criminal confinement as method of punishing mentally ill as socially undesirable).

¹¹⁰ *Id.* at 79.

¹¹¹ *Id.* at 78.

¹¹² See PELKA *supra* note 2, at 48-49 (discussing early critiques of asylum system).

case, minimized unhappiness).¹¹³ That it is right to involuntarily confine a class of persons because doing so is convenient seems utterly immoral. Surely, no one would argue that someone who robs a rich politician who was on his way to pay a hit man acted morally in robbing the politician, but this seems to be the logical conclusion of Negative Utilitarianism. If aggregate happiness can be achieved by unjust motives, and justice seems in some way linked to morality, this aspect of Negative Utilitarianism must be addressed if we seek to have a truly defensible ethical theory for involuntary commitment.

B. Kantian Ethics and *Parens Patriae*

To say the state, as *parens patriae*, is justified in confining a person against his or her will because he or she is *mentally ill* and *a danger to self*, assumes that it is ethical to deprive a person of her liberty if doing so would benefit that individual. While such a notion is categorically at odds with Mill's Principle of Utility,¹¹⁴ it finds support in Kantian ethics.¹¹⁵

Though Kant himself did not address the issue of civil commitment, his method for dealing with ethical issues still applies. Kant's framework begins—and, as we shall see, ends—with the “Categorical Imperative,” a standard consisting of four formulations.¹¹⁶ Our focus is on the second formulation, the “Formula of Humanity”: “Act in such a way that you treat humanity, whether in your own person or in the person of another, always at the same time as an end and never simply as a means.”¹¹⁷

Kant's formulation addresses how individuals should act towards one another. It requires both “action” to treat others as “ends in themselves,” and “inaction,” to never use another individual “simply as a means.”¹¹⁸ Since Kant emphasizes the *treatment* of others as opposed to how one is affected by such treatment, he focuses

¹¹³ See ARRIGO, *supra* note 25, at 78. See also JOSHUA DRESSLER, CASES AND MATERIALS ON CRIMINAL LAW 33-34 (5th ed. 2009) (explaining general idea of laws is to increase happiness in community).

¹¹⁴ See *supra* note 76 and accompanying text (stating Mill's Principle of Utility).

¹¹⁵ See Norman Wilde, *Kant's Relation to Utilitarianism*, 3 PHIL. REV. 289, 302-04 (1894) (describing the relationship between utility and Kantian ethics).

¹¹⁶ See Immanuel Kant, *Groundwork for the Metaphysic of Morals*, in CLASSICS OF WESTERN PHILOSOPHY 1009-58 (Steven M. Cahn ed., Hackett Publishing Co. 3d ed. 1990). Kant's four formulations are: (1) The Formula of Universal Law of Nature: “[A]ct only according to that maxim whereby you can at the same time will that it should become a universal law;” (2) The Humanity Formula; (3) The Formula of Autonomy; and (4) The Formula of the Kingdom of Ends. *Id.*

¹¹⁷ *Id.* at 1036.

¹¹⁸ See William Nelson, *Kant's Formula of Humanity*, 117 MIND 85, 87-94 (2008) (explaining the humanity formula); Thomas Hill, *Humanity as an End in Itself*, 91 INT'L J. ETHICS 84, 84-87 (1980) (interpreting Kant's humanity formula).

more on the intent of the actor than the consequence of his or her actions.¹¹⁹ This is in contrast to Negative Utilitarianism which, as we have seen, emphasizes results.¹²⁰

Onora O'Neill, a leading authority on Kantian ethics, asserts that to use someone "simply as a means" or as "mere means"¹²¹ is to involve them in an action in which they could not in principle consent.¹²² Moral philosopher Christine Korsgaard notes, for example, that if I ask you to lend me money, but have no intention of paying you back, then I am tricking you into contributing to an end (gaining wealth) where you had no opportunity to choose the resultant action.¹²³ She argues that under Kant's "Formula of Humanity," "neither my way of acting nor the end produced by it are things that you are in a position to accept or reject, and this renders them morally wrong."¹²⁴ Likewise, in the case of coercion, O'Neill illustrates that, "[i]f a rich or powerful person threatens a debtor with bankruptcy unless he or she joins in some scheme, then the creditor's intention is to coerce; and the debtor, if coerced, cannot consent to his or her part in the creditor's scheme."¹²⁵

Both actions, these scholars argue, are ethically wrong under Kant's formula.¹²⁶ In each case the actor uses another solely for his or her own benefit, without either the knowledge or consent of the other party. Moreover, these cases are distinguishable from those in which both parties tacitly agree to the action. For example, when you cash a check at the local bank, you use the bank teller—who you need to retrieve your funds—as a means. The teller, likewise, uses you as a means to earn his or her living. In this scenario each party *consents* to his or her part in the transaction.¹²⁷ In the same way, Kant distinguishes between using another as a means, and as a *mere* means.

More complicated, however, is understanding how to act towards others as "ends in themselves."¹²⁸ To begin, Kant views humanity as an ultimate end in itself.¹²⁹ This stems from his view of humans as rational agents who are both free and autonomous.¹³⁰ Under Kant's view, rationality includes the capacity and disposition to act on principles or maxims—a necessary requirement if Kant's theory is to have any

¹¹⁹ See Onora O'Neill, *A Simplified Account of Kantian Ethics*, in MATTERS OF LIFE AND DEATH 411, 413-14 (Tom Regan ed., McGraw-Hill Publishing Co., 1986) (1985), available at <https://users.manchester.edu/Facstaff/SSNaragon/Online/texts/201/O'Neill,%20Kant.pdf>.

¹²⁰ See *supra* notes 77-78 and accompanying text.

¹²¹ See O'Neill, *supra* note 119, at 412. Kant uses both phrases simultaneously. See, Kant, *supra* note 116, at 1039 ("[A]ll rational beings stand under the law that each of them should treat himself and all others never *merely as means* but always at the same time as an end in himself.") (emphasis added).

¹²² See O'Neill, *supra* note 119, at 412.

¹²³ Christine M. Korsgaard, *Creating the Kingdom of Ends: Reciprocity and Responsibility in Personal Relations*, 6 PHIL. PERSPECTIVES 305, 309 (1992) (exploring why people hold others accountable).

¹²⁴ *Id.*

¹²⁵ O'Neill, *supra* note 119, at 412.

¹²⁶ See Korsgaard, *supra* note 123, at 309; see also O'Neill, *supra* note 119, at 412.

¹²⁷ See O'Neill, *supra* note 119, at 412.

¹²⁸ See Hill, *supra* note 118, at 87-89 (constructing several definitions of "ends in themselves" theory from Kant's earlier writings).

¹²⁹ See *id.*

¹³⁰ See e.g., *id.*

effect.¹³¹ In addition, he acknowledges the unique cognitive abilities humans possess, such as abstract thinking, which other beings, such as wild animals, lack.¹³² All this leads Kant to place a high value on humanity and, like Bentham and Mill, on individual liberty as well.¹³³

Given the faith Kant places in humanity, there is both a positive and negative way of reading the imperative to treat others as ends in themselves. In the positive sense, Korsgaard posits that “to treat another as an end in itself is to make her ends your own,” echoing Kant’s reasoning in *Grounding for the Metaphysics of Morals* that “[f]or the ends of any subject who is an end in himself must as far as possible be my ends also, if that conception of an end in itself is to have its full effect in me.”¹³⁴ In this vein, O’Neill asserts that to treat others as ends in themselves is to act towards them in a way that takes into account their own goals or objectives—i.e., their own maxims as rational agents.¹³⁵ Put another way, it means to value their humanity and treat their unique desires as relevant factors in determining how to act towards them.

In the negative sense, treating another as an end in herself means that you should never treat her as a “mere means in an end.”¹³⁶ Put another way, it means to value their humanity and treat their unique desires as relevant factors in determining how to act towards them.¹³⁷ If, for example, a doctor diagnoses her patient incorrectly in order to charge more money, the doctor acts fraudulently or at least dishonestly, using the patient as a mere means to her economic ends. If, however, the doctor takes into account the patient’s humanity, she will likely act towards him in a way that provides the most accurate diagnosis, leaving both parties satisfied.¹³⁸

Korsgaard adds that to treat another as an end in herself is to “respect her autonomy—to leave her actions, decisions, and ends to her own choice.”¹³⁹ Thomas

¹³¹ *Id.* at 86 (stating that humanity can be attributed to even the most foolish and depraved persons).

¹³² *Id.* at 86 (“[H]umanity is thought to include a kind of freedom which lower animals lack—ability to foresee future consequences, adopt long-range goals, resist immediate temptation, and even to commit oneself to ends for which one has no sensuous desire.”).

¹³³ *Id.* at 98.

¹³⁴ Korsgaard, *supra* note 123, at 309.

¹³⁵ *Id.*

¹³⁶ O’Neill, *supra* note 119, at 412.

¹³⁷ *Id.* “To treat someone as an end in him or herself requires in the first place that one not use him or her as mere means, that one respect each as a rational person with his or her own maxims.” *Id.* at 412.

¹³⁸ *See id.* at 412-13 (asserting beneficence requires some maxims must be enacted upon to foster others’ ends). O’Neill reads this, however, as a discretionary component of Kant’s maxim, arguing that it would be difficult if not impossible always to act in order to foster another’s goals, particularly where they may be inconsistent with your own set of maxims. *Id.* That issue, however, is beside the point. Because we have construed “Negative Utilitarianism” as a compliment to Kantian ethics, our focus on Kant’s second maxim is likewise limited to its negative aspect, namely, the imperative not to use someone as “mere means.” *Id.*

¹³⁹ *See* Korsgaard, *supra* note 123, at 309. Korsgaard states:

In the positive sense, to treat another as an end in itself is to make her ends your own: ‘for the ends of any subject who is an end in himself must as far as

Hill, however, argues that while liberty and autonomy have a “high priority” under the Formula of Humanity, it is limited by a concern for the liberty and rational development of all.¹⁴⁰ While Hill does not develop this argument further, he nevertheless suggests the possibility of justifying limitations on personal autonomy and liberty under a Kantian framework.¹⁴¹

Each of the four formulations for the Categorical Imperative bolster one’s argument for finding an action ethical, though all may not apply in a particular case. Because our ethical theory must apply to the state, it must apply uniformly. Consequently, formulations which can guide actions on a large scale are especially helpful in seeking to justify the state’s police power. Kant’s Formulation of Universal Law, for example, is instructive: “Act only according to that maxim by which you can at the same time will that it should become a universal law.”¹⁴²

Kant believes rules of law are ethical only insofar as the law-giver can rationally will that they apply in all cases.¹⁴³ Because a law adhering to this formula is broad enough to apply universally (i.e. to all American jurisdictions), and specific enough so that its application does not depend on additional rules, the maxim is self-sustaining. It does not depend, as in the case of utilitarianism, on several maxims. This is why Kant’s ethical theory begins and ends with the Categorical Imperative.

The rationale underlying the state’s involuntary commitment of those who are *mentally ill* and a *danger to themselves* is that the state as *parens patriae* has a duty to act in the general welfare of those who cannot act for themselves. The state has already demonstrated its will to make *parens patriae* doctrine, as applied in the context of involuntary commitment, a universal law. The fact that both case law and individual state statutes authorize involuntary commitment, where a person is *mentally ill* and a *danger to self*, strongly suggests the state’s desire to apply its *parens patriae* power uniformly in the context of civil commitment—at least to this version of the standard for involuntary commitment.¹⁴⁴ The question remains, however, whether the state’s *parens*

possible be my ends also, if that conception of an end in itself is to have its full effect in me.’

Id. (citing Immanuel Kant, *Metaphysics of Morals*, Prussian Academy Edition Volume IV 430 (1785), translated in James Ellington, *Immanuel Kant. Ethical Philosophy*, Indianapolis-Hackett 37 (1983)).

¹⁴⁰ See Hill, *supra* note 118, at 98 (acknowledging limitations imposed on liberty’s high priority status in regards to Kant’s second maxim).

¹⁴¹ *Id.* (noting Hill’s failure to discuss how such limitations should be applied).

¹⁴² See *supra* note 116 and accompanying text; see also Christine Korsgaard, *Kant’s Formula of Universal Law*, 66 PAC. PHIL. Q. 1, 24 (1985).

¹⁴³ See *id.* at 1 (“Kant says that this is equivalent to acting as though your maxim were by your will to become a law of nature, and he uses this latter formulation his examples of how the imperative is to be applied.”); see also Pauline Kleingeld, *A Contradiction of the Right Kind: Convenience Killing and Kant’s Formula of Universal Law*, 69 PHIL. Q. 64, 65-66 (2018) (instructing how to analyze Kant’s Formula of Universal Law).

¹⁴⁴ See Testa & West, *supra* note 7, at 32. The doctrine of *parens patriae*, the government’s obligation to provide for the incapacitated, was the basis of state commitment standards dating back to mid-19th century. *Id.* See also John A. Menninger, *Involuntary Treatment: Hospitalization and*

patriae power, in this context, comports with Kant's second maxim, that we treat others never as *mere means* but rather as *ends in themselves*.

To begin, we have noted that the *parens patriae* doctrine reflects the notion that a State is authorized to act in the interest of individuals who cannot act on their own.¹⁴⁵ If we adopt Kant's view, it seems that as long as the state is discharging its duty as *parens patriae*, while at the same time acting toward an individual as an end in him or herself, then both maxims are satisfied.

The problem here is what Mill recognized in his Harm Principle—that the value of individual liberty is too great to allow the state to restrict it in any case other than to protect others from harm.¹⁴⁶ The state's interest in helping the individual, Mill argued, is never a sufficient reason.¹⁴⁷ However, Mill was concerned with individuals who were otherwise not suffering from a mental illness that posed a threat to others.¹⁴⁸ Mill's value on personal liberty, specifically in decision-making, is therefore greatly diminished in the context of involuntary civil commitment where the individual is determined unable to make rational choices. The risk of harm to self is something the state seeks to prevent but not something the individual can control, at least at the time commitment is deemed necessary. Therefore, because the intent to prevent such harm focuses on the person as an end in themselves, Kant's second maxim seems readily available to support that course of action. While scholarship on this point is severely deficient, bioethicists and psychiatrists have connected the Principle of Humanity to "medical paternalism" as a justification for promoting a person's well-being where they themselves lack the means to do so.¹⁴⁹

There are three additional problems with our ethical theory. First, scholars have argued that Kant's ethical theory becomes problematic when it concerns individuals with diminished capacity such as those with mental illnesses.¹⁵⁰ This is because Kant's entire ethical framework relates to "rational agents."¹⁵¹ Kant values

Medications, in PSYCHIATRIC SECRETS 477, 478 (2nd ed., 2001) (discussing how state's authority to commit individuals stems from legal theory of *parens patriae*).

¹⁴⁵ *Supra* note 9 and accompanying text.

¹⁴⁶ *Supra* note 79 and accompanying text.

¹⁴⁷ *Supra* note 83 and accompanying text.

¹⁴⁸ Hull, *supra* note 79, at 137-38. Hull highlights this point as being in tension with the second version of the standard. *Id.* "A decision that is autonomous, in the sense of being free from coercion, knowledgeable of alternatives and of relevant consequences, and accepting of potential risks." *Id.*

¹⁴⁹ See John S. Callander, *Ethics and Aims in Psychotherapy: A Contribution From Kant*, 24 J. MED. ETHICS 274, 275-77 (1998); Michael Robertson, *An Overview of Psychiatric Ethics*, NSW INST. MED. EDUC. & TRAINING, 12-13, https://www.slhd.nsw.gov.au/services/PTN/Resources/Overview_Psychiatric_Ethic_Handbook.pdf (last visited Mar. 12, 2020).

¹⁵⁰ See, e.g., Patrick Frierson, *Kant on Mental Disorder. Part 2: Philosophical Implications of Kant's Account*, 20 HIST. PSYCHIATRY 290, 305-06 (2009) (noting those with mental illness may lack free will and ability to act ethically).

¹⁵¹ See Alan Thomas, *Is Your Mind Your Brain?* 1 RICH. J. PHIL., 1, 2-4 (2002) (interpreting Kant's second maxim). "Kant's [second] formulations of the categorical imperative—the injunction never to treat *rational agents* merely as means but rather always to treat them as ends in themselves." *Id.*

humanity as a good in and of itself partly because humans typically possess a “rational capacity” not otherwise available to animals or objects.¹⁵² Consequently, when an underlying assumption for confining a person is the fact that she lacks or has a diminished rational capacity, the interest in valuing the humanity of that individual, the argument goes, is non-existent. Ironically, while the issue of diminished capacity may have solved Mill’s liberty problem, it presents a separate Kantian problem that needs to be addressed.

Second, O’Neill noted that using someone as a mere means relates to actions directed towards individuals to which they are unwilling to consent.¹⁵³ If we understand that most individuals involuntarily committed are so *mentally ill* that they are a *danger to themselves*, then the issue is what happens when both (1) consent is not possible, (2) but the action is not considered fraud or misrepresentation?

Finally, just as Negative Utilitarianism presents its own problems for being consequence-based, so does Kantian ethics for being intent-based.¹⁵⁴ Under Kant, as long as you intended to treat someone as an end in themselves and never as a mere means, and you act towards them in that way, the issue of whether any benefit actually occurs is totally overlooked.¹⁵⁵

Having set out the broad parameters of Negative Utilitarianism and Kantian ethics, as well as what appears to be their most significant defects, we move towards understanding how these theories work together. In so doing, we will also further address their respective deficiencies. While it would be foolish to pretend to perfectly counter each of these contentions, at the very least we can attempt to undermine them as they relate to civil commitment.

¹⁵² See Geoffrey Sayre-McCord, *Rational Agency and the Nature of Normative Concepts*, 4-6, (Jan. 23, 2013), <https://philosophy.unc.edu/files/2013/10/Rational-Agency-and-the-Nature-of-Normative-Concepts1.pdf>. “Agents who have [the] capacity to act on the basis of their normative judgments were valorized by Kant as ‘Rational Agents.’” *Id.*

¹⁵³ See O’Neill, *supra* note 119, at 412. “To use someone as a mere means is to involve them in a scheme of action *to which they could not in principle consent.*” *Id.* (emphasis added). O’Neill goes on to explain that “a person who promises falsely treats the acceptor of the promise as a prop or a thing and not as a person,” further validating the revocation of fundamental rights for involuntarily committed persons. *Id.*

¹⁵⁴ See Wilde, *supra* note 115, at 290 (quoting Mill).

¹⁵⁵ See *id.* at 289. Wilde quotes Kant:

A good will is good not because of what it performs or effects, not by its aptness for the attainment of some proposed end, but simply by virtue of the volition, that is, it is good in itself and considered by itself is to be esteemed much higher than all that can be brought about by it in favor of any inclination, nay, even of the sum total of all inclinations.

Id.

V. Conclusion

One of the dangers of assessing legal standards—such as those for involuntary civil commitment—through the lens of philosophy is the risk of suffocating them with abstract critiques which have no practical effect. This speaks to the counter argument for Negative Utilitarianism (i.e., that it focuses solely on consequences) and the third counter argument for Kantian Ethics (i.e., that it focuses solely on intent). Two reasons suffice to show why such concerns do not diminish the strength of each ethical theory as justifications for the parallel standards of civil commitment.

To begin, the issue of focusing too much on either intent or effects assumes that only one of the ethical theories applies to all involuntary civil commitment proceedings. In other words, it would assume, for example, that a state's authority to commit an individual is based solely on its police power, or solely on its power as *parens patriae*. Yet, as noted throughout this paper, both theories apply concurrently in nearly all fifty states and therefore the risk that civil commitment practices will only look to either effects or intent is unwarranted even on a theoretical level. Indeed, if an individual is committed because they are deemed a danger to others, it seems controversial to say that are also a danger to themselves. On a practical level, the trajectory of disability rights in the United States also works against this position. At its inception, civil commitment was originally based on the state's police power and then almost uniformly shifted to a rationale premised under *parens patriae* logic—at no point was there an equal balance. The current dual-standard we see today, however, represents a conscious policy choice by the government to reach a middle ground in utilizing its binary sources of authority in a way that does not allow either source, the standards they enable, or the ethical principles underlying them, to dominate.

Other concerns, dealing primarily with Kantian ethics, are equally unproblematic. There are several approaches to the second concern, for example, which deals with a patient's inability to consent. One view is simply to say that an individual, by virtue of choosing to reside within the United States, has given tacit consent to being subject to all laws within its borders not the least of which includes civil commitment laws. Another view is that the state is acting consistent with what the will of a confined individual would have been had he or she been competent. This is evidenced by the fact that such individuals are routinely released from civil commitment once they regain coherence and are otherwise able to reside freely in safety. This "patient's best interests" approach assumes that most people, when unaffected by a mental illness, are likely to desire a full recovery if they should ever be so adversely affected by a mental illness as to require involuntary commitment in the first place.

The final view argues that Kantian ethics values humans as ends in themselves only because humans have the capacity for rational thought. It follows, the argument goes, that where such capacity is absent then the value warranting favorable treatment of human beings ceases to exist (i.e., that the state should utilize its *parens patriae* power to act in an individual's best interests). This position, however, illustrates the short-sighted view that mental illness is a fixed and static condition. First, many individuals under civil commitment have the potential for future rationale capacity, indeed that is often the basis for treating them in the first place. This potential alone should suffice for the state to find value in acting towards such individual as ends in themselves. Second, even during admission, many patients have varying levels of lucidity and recognizance, all

which undermines the notion that all persons committed are totally and irrefragably incompetent and thus unworthy of respect and dignity. Such a conclusion is simply inconsistent with how mental illness functions.

As noted previously, these binary standards and their ethical underpinnings work side by side in a continuing attempt by the state to recognize both the humanity of each individual while at the same time honoring its obligation to protect society in general. Even with their defects, the two current standards for involuntary civil commitment together function as an ethically defensible means of satisfying that dual obligation. Thus, while not perfect, we can at the very least say the current standards present a clear advancement in mental health law over all the policies previously utilized.