

A Genealogy of Autonomy: Freedom, Paternalism, and the Future of the Doctor–Patient Relationship

QUENTIN I. T. GENUIS*

The University of British Columbia, Vancouver, British Columbia, Canada

*Address correspondence to: Quentin I. T. Genuis, MD, The University of British Columbia, Department of Emergency Medicine, Diamond Health Care Centre 11th floor, 2775 Laurel Street, Vancouver, BC Canada V5Z 1M9. E-mail: qgenuis@ualberta.ca

Although the principle of respect for personal autonomy has been the subject of debate for almost 40 years, the conversation has often suffered from lack of clarity regarding the philosophical traditions underlying this principle. In this article, I trace a genealogy of autonomy, first contrasting Kant's autonomy as moral obligation and Mill's teleological political liberty. I then show development from Mill's concept to Beauchamp and Childress' principle and to Julian Savulescu's non-teleological autonomy sketch. I argue that, although the reach for a new principle to guide choices in physician–patient relationships can rightfully be seen as important, the notion that is now called autonomy within bioethics has corollaries that undermine critical aspects of medical care. As such, there is need for a richer account of the interplay between the free choice of patients and the informed recommendations of doctors.

Keywords: *autonomy, bioethics, Mill, paternalism, Savulescu*

I. INTRODUCTION

I first learned about the significance of autonomy while I was in medical school, but it was not something I learned in the classroom. I recognized its importance while moonlighting as a leader for a community youth group, particularly when I was involved in counseling teenagers who were considering self-harm. Although these roles—physician-in-training and youth worker—may appear to have little in common, the similarities run strikingly deep. In each I spent a great deal of time on relatively uninteresting

preparatory tasks—memorizing nerve plexuses, practicing physical-exam skills, and writing hospital orders by day; and preparing lesson-plans, organizing leaders, and pre-ordering pizza by night—that paved the way for the engagement that really mattered. In both settings, what really mattered was surprisingly similar: sitting in a room with a vulnerable person for whom I cared; offering advice and education based on my acquired knowledge, expertise, and experience; and hoping somehow to help the person to whom I was talking.

What truly surprised me were the similarities of the conversations across these two settings. In both, there comes a moment when physician and youth worker alike are trained to ask an open question, something like: “Is there anything else you would like to talk to me about?” Often, if the patient or student perceived genuine care behind the open question, this moment paved the way for a deeper conversation. The topics were the same: loneliness and relationships, apathy and purpose, worry and trust, sadness and hope, and brokenness and healing. And, in both situations, my response was the same: provide a listening ear while gently seeking to point toward what I saw as good ends for the patient or student. I could almost-interchangeably refer to my hoped-for ends as *health*—if we are using a World Health Organization definition: “a state of complete physical, mental and social well-being and not merely the absence of . . . infirmity” ([Preamble to the Constitution of the World Health Organization, 1948](#))—or as *eudaimonia*: total prosperity, happiness, blessedness. It has often been remarked that the doctor takes on a priestly role in the present era. We shall see that there are pitfalls to be avoided in such a conception, but it is certainly true to my experience that there is considerable overlap to the roles of the modern doctor and pastor in terms of the “intimate” and even “spiritual” issues they are asked to address ([Pink, Jacobson, and Pritchard, 2007](#), 841).

It was during this time of both medical and “pastoral” engagement that I was introduced to a particular history-of-medicine narrative that cast doubts on my approach in both contexts. A simple sketch of the narrative follows: Western medicine long operated in a Hippocratic tradition, the chief principle of which suggests that the role of the doctor is to “use his or her judgment to try to benefit the patient and protect the patient from harm” ([Veatch, 1989](#), 47). Although this approach allowed for “remarkable continuity in medical ethics across millennia,” the advent of pluralistic society revealed a deeply questionable—even unethical—tendency within it: to impose one particular perceived good on a patient who may have an entirely different conception of the good or who may simply weigh competing goods differently ([Beauchamp and Childress, 2013](#), 1). Such action is viewed as paternalistic and as fundamentally disrespecting each individual’s right to self-determination. Given these problems, medicine in the twentieth century left the Hippocratic tradition behind and sought a new principle that might avoid the pitfall of paternalism. This project reached its zenith in [1979](#), when

bioethicists Tom Beauchamp and James Childress published *Principles of Biomedical Ethics*, a work that identified personal autonomy—a term with considerable import in philosophy since Kant—as a notion that would provide the way forward. Introducing respect for autonomy as a principle of bioethics reoriented patient care around value-neutral education of the patient, such that he or she might make un-coerced and informed—and therefore truly free—decisions that align with his or her chosen goals.

Although there is considerable nuance to the story of what we now call bioethics that is not captured in this narrative, it does successfully point to a key shift that took place with the introduction of a strong autonomy principle into the conversation. Furthermore, this account seems to make clear how a respect for autonomy principle provides a way forward for medicine within a pluralistic context. If we cannot agree on what is right and we want to safeguard against manipulation, we can at least enable individuals to seek their own goods in their own ways. Sitting in an ethics lecture, this seemed to make good sense.

Now it made less sense in the setting of actual encounters with people for whom I cared. This was particularly true in the heartrending cases of teenagers contemplating self-harm. Such cases saddened me both in the startling frequency with which I encountered them and in the disconcerting lucidity with which many people arrived at such a desire. By almost any contemporary measure of autonomy, most young people I encountered in these cases would be deemed autonomous. If some of my ethics professors were to be believed, the appropriate response for a medical professional in such cases might look something like this: upon hearing that an autonomous individual is considering self-harm, I should educate the individual as to the potential ways forward. Wanting to avoid manipulation, I would cautiously avoid statements that might be rooted in my worldview, emotional reaction to the situation, spiritual perspective, or personal opinion regarding the best way forward. Once satisfied that the individual has been provided with sufficient information to make a well-educated choice, I would leave that individual to make an autonomous decision.

I, however, did not approach such situations in this fashion, for it seemed obvious that to do so would be to neglect compassion—to abandon a suffering individual whom I was committed to helping. Instead, I worked to convince those who were considering self-harm that their lives were incredibly valuable, contrary to their own perception. I deliberately used statements and rhetorical techniques that emphasized my genuine concern for them. I refused to abandon them to the terror of their own self-rule. Most basically, my stance toward these individuals could easily be labeled as paternalistic: viewing the autonomous desire of another individual as unhealthy and seeking to steer them toward what I saw as a better path.

Yet, I would not readily welcome being described as a paternalistic caregiver. This is true not only because of the negative stigma associated with the

term in the context of the bioethics dialogue. I would also resist such a characterization because I view what the autonomy movement has imprecisely called paternalistic orientation—the tendency within medicine to confuse medical expertise with more comprehensive authority to make moral, ethical, ideological, and social judgments for patients—as genuinely dangerous.

Another brief, and inevitably coarse, history-of-medicine sketch may be instructive here. For much of history, doctors in Western medical traditions had a chiefly medical role in their patients' lives: aiming to diagnose and treat illness, and not attempting to occupy the "biopsychosocial-spiritual" role that many doctors strive to adopt today (Sulmasy, 2002, 24). In this narrow role, doctors behaved with a certain degree of "benign paternalism" in the clinical decision-making process (Pellegrino and Thomasma, 1993, 127). As noted above, modernity invited the physician to play a priestly role in the lives of patients. Thus, a doctor's role morphed from that of a medical authority to one who seeks the patient's general well-being: a holistic state that includes social, spiritual, and moral aspects. The doctor's influence in such a setting had the potential to become almost totalizing; in this bloated role, medicine was tempted toward imperiousness and peremptoriness. Sometimes, it succumbed, which prompted the need for patient-centered reform. In such a scenario, it is easy to see how limitations on the paternalism of the doctor, via an autonomy principle, might seem necessary to prevent manipulative or overreaching uses of the medical profession's new-found influence. If physician paternalism connotes coercion or reaching out of the medical sphere to seek totalizing or controlling influence over another person, it is necessary to find a way beyond it. This is to say that, although I struggled to reconcile the autonomy principle that I had been taught with good care, I empathized with the pursuit of a "non-paternalistic" account of the roles of doctor and patient in medical and ethical decision making. Now why is the principle of respect for personal autonomy unable to provide such an account? And, if it cannot, what might?

This article comes out of these questions and seeks to begin to address them. To that end, I trace a genealogy of autonomy language to show how the dominant contemporary understanding of the term not only differs from the earlier philosophies on which it purports to rely, but also has unintended corollaries that undermine important aspects of medical care. In order to provide groundwork for this argument, it is first necessary to survey Kant's theory of autonomy as moral obligation and Mill's teleological political liberty. These sketches reveal how, as has been discussed elsewhere,¹ the bioethical principle of respect for personal autonomy that was first proposed by Beauchamp and Childress can be seen as a development of Mill's thought but not that of Kant. Then it is possible to show how some contemporary thinkers, notably Julian Savulescu, recognize Millian liberty as the forerunner of bioethical autonomy and purport to embrace a Millian framework, while ironically neglecting Mill's own fundamental convictions. Clear tracing of this

genealogy is required if we are to understand 1) the difficulties attending Savulescu's *non-teleological* approach to autonomy and 2) the critical need for a richer account of the interplay between the free choice of a patient and the informed recommendations of a physician in a clinical context.

II. AUTONOMY BEFORE BIOETHICS

In this section, I outline relevant philosophies of autonomy prior to the importation of the term into bioethics, for the purpose of later showing how these earlier ideas have been used, and misused, by more contemporary thinkers. After briefly outlining the way *autonomia* was used in the Greek tradition, I contrast Kantian autonomy with Millian liberty. In comparing these two views, I specifically note their implications regarding self-harm and suicide, issues that are explicitly discussed in Kant and Mill and are manifestly relevant to contemporary bioethics. This analysis lays the groundwork for evaluating the success of Beauchamp and Childress' attempt to "harmonize Kant and Mill on autonomy," and the cogency of Savulescu's understanding and use of Mill (Lysaught, 2004, 675).

First uses

It is well known that the English "autonomy" is a transliteration of the Greek *autonomia* (*autos*—"self" and *nomos*—"rule," "law," or "governance"). There is, however, no meaningful philosophical continuity between the ancient Greek *autonomia* and the principle of respect for individual autonomy as it exists in contemporary bioethics. *Autonomia* did not have any individual ethical meaning in antiquity. Rather, in ancient writings, *autonomia* is almost exclusively a political term, relating to the self-governance of a given community, usually a city-state. It was not until Kant's *Grounding for the Metaphysics of Morals* was published in 1785 that the term "autonomy" began to be understood within a lexicon other than that of "inter-state relations" (Otswald, 1982, 1). As such, the present survey commences with the function of autonomy within Kant's moral thought.

Autonomy in Kant's thought

Grounding for the Metaphysics of Morals clearly states Kant's general definition of autonomy: "the property that the will has of being a law to itself" (1993, 440). Given Kant's belief in the universality of moral law, a more specific definition follows: "man is subject only to his own, yet universal, legislation . . . he is bound only to act in accordance with his own will, which is, however, a will purposed by nature to legislate universal laws" (1993, 432). How might such a conception of autonomy work, and how is it relevant to bioethics?

It is initially important to note that Kant's autonomy is a positive notion. In other words, autonomous action is not simply that which is undertaken in the absence of controlling interference; rather, a truly autonomous act is a rational and actively chosen one that conforms to a principle of universality. Actions are autonomous only if they comply with Kant's a priori categorical imperative: "[a]ct only according to that maxim whereby you can at the same time will that it should become a universal law" (1993, 421). As such, autonomy, according to Kant, is a principle of moral philosophy that guides the rational individual toward universalizable duties or obligations.

This conception relates to Kant's underlying theory of morality. In *Religion within the Bounds of Bare Reason*, he assumes that "[m]orality . . . is based on the concept of the human being as one who is free, but who precisely therefore also binds himself through his reason to unconditional laws" (2009, 1). Because human beings, to cite another of Kant's famous aphorisms, "exist as ends in themselves," there is significant moral value present in a willful and dutiful good act that is not present in that same good act when it is dictated by an object external to the self (1993, 428). Herein lies Kant's opposition to what he refers to as heteronomy. The autonomy–heteronomy distinction in *Grounding for the Metaphysics of Morals* presents an imperative to act based on the categorical imperative, not on desires or a posteriori (or empirical) cognition. Kant commentator Robert Paul Wolff summarizes clearly: Kant "makes it clear that by heteronomy he means bondage to *objects* outside the self, rather than bondage to other *wills*. The bondage may arise out of desire, as for pleasure or happiness, or it may arise out of a rational conception, such as the principle of perfection" (1986, 187). Actions that arise out of this type of bondage is unacceptable to Kant: "the proposition, The human being is *evil*, can signify nothing other than this: He is conscious of the moral law and yet has admitted the (occasional) deviation from it into his maxim" (2009, 32).

Kant's notion of autonomy is, therefore, extremely difficult to reconcile with some recent portrayals of "Kantian autonomy" within the field of bioethics. For example, Thomas Mappes and David DeGrazia's popular textbook *Biomedical Ethics* characterizes Kantian subjects as those who legislate "their own actions in accordance with rules of their own choosing" (1996, 28). They go on to say that a "Kantian position central in biomedical ethics describes autonomy in terms of self-control, self-direction, or self-governance" (Mappes and DeGrazia, 1996, 28). While this is not technically incorrect, it misleadingly neglects the moral content of Kant's autonomy: that actions in accordance with rules of one's own choosing are *only* autonomous if the chosen rules are universal moral laws. Barbara Secker effectively summarizes this contention: "Kant's conception is not of individual or personal autonomy, where the central question is "What do I really want, and is it best for me?" rather, it is of moral autonomy which applies universally, and asks the question "Is this what I ought to do?" morally speaking"

(1999, 48). As such, Secker contends that it is useful to distinguish sharply between “Kant’s autonomy” and “Kantian autonomy”—the former being a moral notion based in Kant’s own thought and the latter as a bioethical principle that originates elsewhere. Secker contends that these two notions have “very little at all to do with” each other (1999, 43–4). While I concur with this logical analysis, such terminology undoubtedly lends itself to confusion. If autonomy principles in bioethics do not function congruously with Kant’s moral understanding of the term, it seems important to look elsewhere to discern where they come from. What Secker is referring to as Kantian autonomy—the concept that many bioethicists simply call autonomy—is actually a development of Millian liberty. We turn shortly to examination of this notion.

What might Kantian autonomy say to bioethics? Perhaps a great deal, but not in the ways that autonomy language is usually used. Kant’s autonomy–heteronomy distinction is wholly different from the autonomy–paternalism tension often discussed in bioethics. *Grounding for the Metaphysics of Morals* is silent on the issue now referred to as paternalism. Kant does, however, mention the implications of his conception for first-party harm: “To preserve one’s life is a duty” (1993, 397). The logic is simple: to will one’s own harm would violate the categorical imperative; realizing this truth enlightens the mind to a duty that is morally binding. For Kant, there could be no such thing as an autonomous request for euthanasia or Physician-Assisted-Death (PAD). It is vital to keep our categories straight here. Kant is not necessarily suggesting that suicide, euthanasia, or PAD should be illegal; nor does he speak to a question of the extent to which doctors should facilitate the nonautonomous wishes of their patients. Kant simply asserts that desiring to die cannot be autonomous and is therefore immoral. It is easy to see how much of contemporary autonomy thought represents a radical departure from such an understanding.

Mill’s liberty

It has been antecedently noted that, although Mill never uses the term autonomy, his understanding of human freedom has had profound impact on Western culture, including bioethics. His famous understanding of the concept in *On Liberty* (1859) follows: “The only freedom which deserves the name is that of pursuing our own good in our own way, so long as we do not attempt to deprive others of theirs or impede their efforts to obtain it” (Mill, 2005, 15). It should be noted that Mill’s freedom is a teleological notion, that it has implications with regard to self-harm, and that it is a *political* formulation.

The logic of *On Liberty* is relatively straightforward. Recognizing himself to be living in a plural world where unanimous agreement on societal or personal goods is likely to be impossible, Mill asserts his definition of freedom as a way by which diverse peoples can live at peace with one another.

According to this formulation, which Isaiah Berlin has more recently called “negative liberty,” free action is not found through adherence to some duty or principle, as in Kant (1969, xvli). Rather, a free action is simply that which is undertaken in the absence of controlling influence. In order to maximize this liberty for all members of a community, Mill asserts, “the 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” (2005, 13). This vision presents a kind of lowest-common-denominator political ethic that purports to allow “civilized” societies without common cultural or moral norms to function without succumbing to the “tyranny of the majority” (2005, 7).

Mill follows this argument by asserting that first-party harm does not provide “sufficient warrant” for intervening in the decisions of another person (2005, 13). It is worthwhile to briefly consider the logic that seems to be behind Mill’s distinction between first- and third-party harm as grounds for restriction of liberty. As might be expected in the argument of a utilitarian, Mill refuses to posit a simple right to choice. It is not immediately obvious, however, how a simply utilitarian approach would prohibit restraining individuals who were attempting to kill themselves. Here it is crucial to note that, for Mill, “utility in the highest sense [is] grounded on the permanent interests of man as a progressive being” (2005, 14). Conscious choice is seen as a central feature of being human *only because* it protects this permanent interest: the progress of individuals and societies toward higher goods. As such, Mill’s liberty is teleological: not a moral good *in itself* but a necessary political good for societies capable of harnessing choice toward greater human thriving. In such societies, protection of the ability of living persons to exercise choice is the only factor that can allow for restriction of liberty. I here note Mill’s prohibition against selling oneself into slavery: “The principle of freedom cannot require that he should be free not to be free. It is not freedom to be allowed to alienate his freedom” (2005, 125). In Mill’s vision, net utility is lost by allowing individuals to sell themselves as a slave but gained by allowing individuals to commit suicide. The human being may choose to die, but may not choose to live without political liberty. More comprehensive critical evaluation of Mill might question his underlying conception of the human being or carefully examine whether his utilitarianism is grounded in a notion that functions indistinguishably from a quintessentially nonutilitarian right to choice. For the purposes of this analysis, however, it is sufficient to note how Mill’s liberty renders human desires as worthy of near-absolute *political* protection whilst remaining teleological. For Mill, liberty is vital because it provides the necessary conditions from which human thriving may arise. However, liberty is not, in itself, the *summum bonum*.

As with Kant, Mill’s thought is less *directly* relevant to medical ethics than a perusal of most bioethics textbooks might suggest. *On Liberty* begins with the following: “The subject of this essay is not the so-called ‘liberty of the will’ . . . but civil, or social liberty: the nature and limits of the power which

can be legitimately exercised by society over the individual” (2005, 3). Mill is making a narrow political argument regarding what individuals may or may not be allowed to do by governments. To again use euthanasia and PAD as an example: there are important logical gaps between the assertions “the government should not make suicide illegal” and “it is ethically acceptable for a doctor to grant the request of a patient for assistance in dying.” A Millian argument—represented by the former assertion—cannot be dropped into bioethics debates without considering (or unthinkingly assuming certain answers to) the following questions: Is there a difference between what should be legal and what constitutes moral, or ethically correct, action? Is there an ethical difference between legally allowing an action because intervening to stop it might cause greater harm (or loss of utility) and actively assisting that action? And, I think most centrally: does a human person, in his or her specific role as a doctor, relate to a patient in the same way that a governing body relates to individuals within society?

III. BEAUCHAMP AND CHILDRESS

To this point, I have noted Kant’s moral autonomy and Mill’s teleological political liberty. We have noted that, while neither speaks directly to bioethics, both lines of thought might have bioethical implications, most obviously on cases of first-party harm. It is now possible to consider, focusing specifically on Beauchamp and Childress’ *Principles of Biomedical Ethics*, the contention that their principle of respect for personal autonomy is essentially a reappropriation of Millian liberty for a medical context where the doctor is positioned as a value-neutral educator and technician. Later, I also show the beginnings of a divergence from Mill toward the nonteleological autonomy that further develops in Savulescu’s work.

It is well known that Beauchamp and Childress propose four basic principles that they see as capable of guiding proper biomedical ethics: respect for autonomy, beneficence, non-maleficence, and justice. It is also well known that their “principle of respect for the autonomous choices of persons” has been met with heavy criticism (Beauchamp and Childress, 2013, 101). Commentators have argued that this conception of autonomy is ultimately arbitrary; that it is “self-defeating” (Pellegrino and Thomasma, 1993, 193); that it positions the physician–patient relationship in such a way that precludes trust and compassion; that it is “not valuable enough to offset what we lose by leaving people to their own autonomous choices” (Conly, 2012, 1); that it is excessively individualistic and thus produces an ethic of selfishness and self-indulgence; that it produces an “intolerance of dependence on others” (McCormick, 2012, 1083); that it flies “in the face of reality” (Meilaender, 2013, 59); that it “tend[s] to moral atomism and moral anarchy” (Pellegrino and Thomasma, 1993, 193); that it is “only for those in power”

(Keenan, 2012, 1085); and, most centrally, that it is totalizing, subordinating all other principles or considerations. Beauchamp and Childress deny the cogency of such critiques more forcefully with each successive edition of *Principles of Biomedical Ethics*.

It seems to me that there are confusions that attend many of these critiques. They often contribute to confusion in the dialogue by aiming at autonomy when they are really trying to oppose relativistic moral individualism—which is not *necessarily* a product of Beauchamp and Childress' thought. Nevertheless, this host of criticisms points us toward a fundamental problem with their principle of respect for autonomy. This problem is valid *if* careful comparison reveals Beauchamp and Childress' principle as essentially Millian. We have seen that, for Mill, liberty is the category by which the goods of human life are achieved. In this view, justice is simply that which ensures maximum liberty; non-maleficence means respect for liberty, even over life; and beneficence is either rendered relative or moot, for absolute beneficence requires reference to an objective good, something Mill aims to protect the individual from. Libertarian autonomy is the “object of obligation” that subordinates and relativizes other ethical principles and obligations (Jennings, 2009, 83). As such, *if* Beauchamp and Childress' conception of autonomy proves to be a bioethical reappropriation of Millian liberty, *then* we might correctly see it as tending toward absolutization.

Given the established contrast between Kantian and Millian understandings of autonomy, it proves relatively easy to show that Beauchamp and Childress' autonomy is essentially Millian, albeit with two additional features that seem to outline the specific role of the doctor: a positive criteria of education and a professional obligation to assist others in actualizing their purportedly autonomous choices. This gives rise to a conception of the doctor as a value-neutral educator and technician. Beauchamp and Childress' basic—and transparently Millian—definition of autonomy follows: “self-rule that is free from both controlling interference by others and limitations that prevent meaningful choice, such as inadequate understanding” (2013, 101). To understand the implications of Beauchamp and Childress' autonomy principle, it is important to examine their “three condition theory,” defense against their critics, treatment of Kant, and approach to first-party harm (2013, 104).

First, Beauchamp and Childress' “three condition theory” defines autonomy by one negative criterion—“noncontrol”—and two positive ones—understanding and intentionality (2013, 104). *Prima facie*, the addition of positive criteria may seem to be a significant departure from Mill. However, Mill believes both that properly free human acts are intentional and that human beings can only properly exercise liberty when they have achieved rational understanding. “Liberty, as a principle, has no application to any state of things anterior to the time when mankind have become capable of being improved by free and equal discussion” (2005, 14). For Mill, this is true both

for societies—as in his defense of what he calls “despotism [as] a legitimate mode of government in dealing with barbarians”—and for individuals—most obviously young children (2005, 14). For Mill, autonomy comes gradually through progress toward rational knowledge. Although this is not a positive criterion for individual political liberty (a government in a “more advanced” society must protect even actions that are not wholly “liberated”), it is easy to see how this is continuous with Beauchamp and Childress’ understanding criterion. In a specialized and highly technologized field such as medicine, it is very difficult for individuals to form self-chosen plans that align with their goals. As such, a central role of the doctor is to provide the information necessary for the individual to progress to a place of autonomous choosing.

Second, Beauchamp and Childress’ defense of their autonomy principle against their critics is quintessentially Millian. They claim that their conception of autonomy is not totalizing because it is subject to the following exceptions: “choices [that] endanger public health, potentially harm innocent others, or require a scarce and unfunded resource” (2013, ix). The first two are obvious cases of third-party harm, and the third, though slightly vague, is either a case of third-party harm (i.e., consumption of public resources to the detriment of others’ care) or utilitarian reasoning (i.e., resource-allocation to produce maximum health at the expense of the health of some). In either case, these examples reveal a Millian conception of autonomy. By contrast, Kant does not view prevention of third-party harm as a restriction of *autonomy* at all.

Third, it is worth briefly noting that Beauchamp and Childress are aware that Kant’s theory of autonomy “differs from” their understanding (2013, 363). However, they see “Kant’s second formulation of the categorical imperative—that persons must be treated as ends and not means only—[as] the substantive basis of the principle of respect for autonomy” that is proposed in *Principles of Biomedical Ethics* (Beauchamp and Childress, 2013, 367). However, this is *only* true, especially in cases of first-party harm, if respecting persons as ends in themselves means respecting their conscious choice, even above their life—a quintessentially Millian view. Thus one may conclude that, even in their attempted re-appropriation of Kantian ethics, these modern thinkers are fundamentally Millian in their approach. As argued above, Kant would not see the second formulation of the categorical imperative as a possible basis for the type of principle that Beauchamp and Childress are proposing.

Finally, as Kant’s and Mill’s respective attitudes toward first-party harm have been antecedently outlined, it is instructive to examine Beauchamp and Childress’ approach to this issue. Here it is helpful to briefly turn to earlier editions of *Principles of Biomedical Ethics*, which explicitly treat the question of autonomy and PAD. By the fourth edition (Beauchamp and Childress, 1994)—and presumably as a response to heavy criticism of the implications of the autonomy principle—all discussion of suicide and PAD

is shifted to other chapters, framing these issues in terms of beneficent and non-maleficent pain relief. This somewhat obscures the issue; it is therefore helpful to draw from the first edition while noting that more recent editions arise out of the same foundational understandings.

The definition of autonomy in the first edition of *Principles of Biomedical Ethics* could easily be a quote from *On Liberty*: “insofar as an autonomous agent’s actions do not infringe [on] the autonomous actions of others, that person should be free to perform whatever action he wishes—even if it involves serious risk for the agent” (Beauchamp and Childress, 1979, 59). Application of this principle to cases of first-party harm reveals the following axiom: “It would . . . be a showing of disrespect to deny autonomous persons the right to commit suicide when, in their considered judgment, they ought to do so” (Beauchamp and Childress, 1979, 87). The task presented to the physician, when responding to a patient who desires to die, is therefore a matter of determining autonomy. Even in the most recent (as of this writing) edition of *Principles of Biomedical Ethics*, this is a process within which “the burden of proof is rightly placed on those who claim that the patient’s judgment is insufficiently autonomous” (Beauchamp and Childress, 2013, 225). In absence of obvious proof that the patient is not acting autonomously—which is to say intentionally and with adequate understanding—the doctor has good reason, perhaps even professional obligation, to assist a patient in committing suicide. Beauchamp and Childress do recognize a kind of felt tension between their autonomy principle and what they call “the principle of human worth” (1979, 87). This second notion, however, is never mentioned again and is given no weight in the ethical decision-making process. As such, to return to treatment of the more recent edition, Beauchamp and Childress commit themselves to a view of autonomy that offers “strong reasons for recognizing the legitimacy of physician-assisted suicide” (2013, 185).

In addition to noting again that this understanding is mutually exclusive with Kant’s autonomy, it is critical to highlight a particular ethical move here. I have shown how Mill’s autonomy cannot be reappropriated for bioethics without assuming answers to certain questions about the physician–patient relationship. Mill does not speak directly to medicine or to the role of a doctor within society. However, in cases of an individual considering self-harm, Mill, although he is the archetypal antipaternalist, sees there being “good reasons for remonstrating with him, or reasoning with him, or persuading him, or entreating him” to refrain (2005, 13). Mill’s argument claims that one must not forcibly prevent suicide, but would have no conflict with my attempts, either as a medical professional or as a youth worker, to dissuade young people from harming themselves. Beauchamp and Childress’ definition of autonomy diverges from Mill in their viewing autonomy as presenting a positive imperative to a doctor to actively facilitate the choices of patients. Here we begin to see transformation of Mill’s *teleological autonomy*—within which choice is politically protected but may still be impelled by other bodies or

individuals—to a *non-teleological autonomy*—where choice itself becomes sacred and may itself compel the actions of third parties. In this ethical move beyond Mill, the second role of the doctor in Beauchamp and Childress' view becomes clear: the servile technician—professionally obliged to assist in the actualization of autonomous wishes.

I therefore conclude that Beauchamp and Childress' principle of respect for personal autonomy 1) is not congruous with Kant's notion of autonomy; 2) arises as a development of Mill's political liberty; and 3) faces immediate problems arising from both the attempt to simply drop political philosophy into bioethics and from some moves beyond Mill and toward a *non-teleological autonomy* that will further develop in Savulescu. A view of the human being, which suggests that "[t]he only evil greater than one's personal death is increasingly taken to be the loss of control of that death," has thus been imported into bioethics (Callahan, 2000, 37). In the name of protecting this control, the doctor is positioned as a value-neutral educator and technician: responsible for ensuring autonomous choosing through teaching and for carrying out the wishes of autonomous patients. Close analysis of the principle of respect for personal autonomy therefore reveals a notion that simultaneously trends toward both totalization—subordinating and relativizing all other ethical principles (including Beauchamp and Childress' own beneficence, non-maleficence, and justice)—and minimalism—reducing the doctor–patient relationship to that of an exchange of goods and services.

It is finally important to note one qualifier to this analysis of Beauchamp and Childress: the above explicates logical implications—not the explicit intentions—of their argument. This is to say that these thinkers are not seeking to totalize autonomy, nor would they celebrate the reduction of the role of a doctor to that of a "value-free . . . set of skills available on the marketplace" (Verhey, 2012, 110). Too few critics of *Principles of Biomedical Ethics* have seen this. It seems to me that Beauchamp and Childress hope that a doctor can be much more than this: a safe confidant, a compassionate caregiver, and perhaps even a trusted advisor with whom the individual-authority dialectic can be transcended through a relationship of mutual respect and trust. Furthermore, Beauchamp and Childress seem to hope that making a *prima facie* principle out of respect for autonomy might encourage patient responsibility and protect against manipulation without tending toward social atomism. These are important goals. But, facing a difficult ethical landscape and seeking a turn away from paternalistic medicine, these authors end up proposing a principle that ultimately subverts these good intentions.

IV. AWAY FROM TELEOLOGY

After Beauchamp and Childress, the autonomy dialogue fragments in predictable ways. On the one hand, the critics noted above have rushed to attack

the principle of respect for personal autonomy. I have noted dissonance in Beauchamp and Childress' account between their intentions for autonomy and the logical and practical implications of their principle. Perhaps dissonance begets dissonance, for criticisms of *Principles of Biomedical Ethics* have often further confused the dialogue. Viewing the autonomy principle as one particular instantiation of some broader issue, critics have often caricatured Beauchamp and Childress' thought. To again cite Jennings' insightful analysis:

Many of these recent criticisms of the effects of the concept of autonomy in our moral discourse and in our social lives cast an overly broad net. Rather than subjecting autonomy to a more careful conceptual analysis, they tend to use the term (or the concept) as a lightning rod for many different social complaints and criticisms. (2009, 80)

This is to say that, oftentimes, autonomy critics who aim at Beauchamp and Childress are *actually* seeking to criticize moral relativism, the so-called "secularization thesis," or—I think most often—what sociologist Alan Wolfe has called "moral freedom," according to which individuals may "defin[e] their own morality" (2001, 199). Because Beauchamp and Childress are not *aiming* at either relativism, moral individualism, aggressive secularism, or an ethic of self-indulgence, they do not seem the best target of such critiques. Here I concur with Jennings, "what is most needed now . . . is not a broad-brush critique but a more careful sifting of the meanings and inconsistencies in the way the term is used" (2009, 80). This more careful project has been, in part, attempted here.

On the other hand, some contemporary scholars have explicitly affirmed notions that I have previously identified as unwanted corollaries of Beauchamp and Childress' formulations. A notable example of this can be found in the work of Julian Savulescu, perhaps most clearly in the 2003 monograph *Medical Ethics and Law: The Core Curriculum*, a work he coauthored with Tony Hope and Judith Hendrick. Where Beauchamp and Childress deny the Millian nature of their framework, this more recent work explicitly affirms Mill as the proper source for theories of autonomy in bioethics. However, as noted above, importation of Mill into biomedical ethics requires making difficult philosophical moves.

Chief among these is the assumption that the space between doctor and patient is essentially a political space, analogous to the space between governing body and individual in Mill's political philosophy. There is significant irony here: the reaction against a supposedly manipulative paternalistic medicine produces a philosophy that cannot view the doctor in any other way than as a potentially coercive governing force, thereby precluding other forms of doctor-patient relationship.

Given an essentially political medical space, the bioethical conversation becomes synonymous with a question of law, as the title of Hope, Savulescu, and Hendrick's monograph suggests. The ethical doctor is simply one who follows the law and helps patients achieve autonomously chosen

ends. Savulescu binds autonomous choices by the categories of capacity and reason, where “capacity is, ultimately, a legal not a medical decision” and reasonable simply means “consistent with the person’s life plans” (Hope, Savulescu, and Hendrick, 2008, 79, 41). As such, Savulescu believes that, in some circumstances, it would be ethically proper for a doctor to provide drugs to addicts or to amputate the leg of an apotemnophilic. Previously, these types of things had only been mentioned as *reductio ad absurdum* critiques of Beauchamp and Childress’s autonomy; here they are explicitly affirmed. With regard to the latter example, Savulescu argues: “doctors should perform surgery, not because it is necessarily in the person’s medical interests, but because it is in his or her overall interests” (2008, 238). Regardless of whether or not one agrees that such a surgery is actually in a person’s overall interests, it is clear that this view renders the task of the doctor as demedicalized. In this view, a doctor’s task is not a task of caring for a patient’s health as much as serving consumer desire. There might be limits to the scope of autonomy, but these are imposed by economics or by law, not by morality or by the scope of medical practice. Savulescu’s treatment of abortion for reasons of gender selection is instructive here. Hope, Savulescu, and Hendrick advise against it in *Medical Ethics and Law* because it is illegal in most of the countries where their book is read, but they see no *ethical* reason to oppose such a choice.

Most foundational among the difficulties attending this importation of Millian liberty into medical ethics is that it actually neglects Mill’s fundamental convictions. As noted above, Mill conceives of liberty as worthy of near-absolute political protection, but only because liberty provides possibility for the progress of mankind. Progress is the ultimate good, but “it is only the cultivation of individuality which produces, or can produce, well-developed human beings” (Mill, 2005, 77). Liberty (or autonomy) is a necessary *political* good here, but it is a *secondary* good. Savulescu’s presentation of Mill blurs this:

The value of individuality for Mill is intrinsic . . . controversial choices are valuable insofar as they promote a better life, a life of more well-being. But they are also *independently valuable* when they are expressions of active decision and deliberation about how to live. There is a value in just deciding to be. (2007, 30)

The focus shifts from autonomy as a negative concept that demands negative reinforcement (i.e., noninterference) because it is necessary for progress, to autonomy as a positive concept that demands positive reinforcement because autonomous choice is a valuable good in itself. If Kant sees autonomy as the only way human beings may realize universalizable goods, and Mill sees it as the only way human societies progress toward good ends, Savulescu sees it as the good in itself: the *sine qua non* of human life.

It is not difficult to see how this approach is a more natural target for critics of autonomy than Beauchamp and Childress are. Savulescu directly suggests

a kind of absolutization of autonomy, and he reduces the role of a doctor to that of a “public servant” whose professional life may not be ethically informed by anything other than law and patient choice (2006, 297). Perhaps counterintuitively, this actually makes Savulescu more difficult to criticize. Because Beauchamp and Childress would likely not support amputating the leg of an apotemnophilic, arguing that such a course of action would potentially be suggested by their ethics proves to be an effective criticism of their proposed framework. Now Savulescu recognizes and affirms these logical corollaries: “[w]e must be open to . . . radical possibilities” (2007, 28). This approach grants Savulescu a kind of logical consistency, but it also quickly eliminates space for common discussion about ethical theory—what factors make a choice right or wrong. Instead, proper medical conduct is determined by political context. Indeed, Tom Beauchamp comments: “This [moral] “theory” part of the landscape of bioethics I expect to vanish soon, because it is serving no useful purpose” (Beauchamp, 2004, 210). In the absence of such a space, there are no grounds on which autonomy advocates and critics might have productive ethical discussion. As such, autonomy discussions within bioethics are becoming increasingly fragmented—lacking common content or narrative—and increasingly shrill—lacking effective means of dialogue.

Most basically, the philosophical stalemate in the autonomy dialogue grows out of competing basic premises regarding human freedom and human will. Either autonomy is bound by ethical norms or it is the object of ethical obligation. Either freedom *is* willing or freedom is teleological: the product of willing in a proper direction. Either Nietzsche or Aristotle. These premises are mutually exclusive, they suggest very different visions for bioethics, and they seem conceptually incommensurable, hence the fragmentation of the moral conversation.²

It seems, to conclude, that the autonomy dialogue is currently in a difficult and ambiguous place. I have here argued that human societies prior to Kant, notably Greek society, did not have any well-developed notion of personal autonomy. In Kant, we see the term introduced into the dialogue of moral philosophy as a way of allowing for rational morality and for seeing individuals as ends in themselves. I have contrasted Kant’s *moral* philosophy with Mill’s *political* philosophy and have argued that the bioethical principle of autonomy finds its root in the latter understanding. I have then shown how some contemporary ethicists have moved beyond Mill’s *teleological* autonomy and have instead embraced *non-teleological autonomy*. I have attempted to show how such a vision “rests on a conflict-ridden and antagonistic picture of social existence” that surely is not a healthy starting place for physician–patient interaction (Jennings, 2009, 85). Other corollaries of these moves include the importation into bioethics of the view that human choice is more ethically valuable than human life; recognition of the space between doctor and patient as an essentially legal or political space; the minimization of the role of the doctor in such a way that potentially precludes care

for patient health; and, ultimately, dissolution of cogent debate about ethics. Each of these consequences seems deeply problematic to me.

As such, the present situation begs for creative response that moves beyond criticism of autonomy language in bioethics. Autonomy language does aim, however misguided, to promote human freedom and responsibility. Furthermore, bioethics undoubtedly needs a lexicon that allows it to value the choices of individual patients as morally significant. As such, though it might be easier to throw rocks at the glasshouses built by post-Enlightenment rationalist thought, constructive thought is needed if we are to move beyond the difficulties of the present situation. It is vital to seek creative reimagination of the discourse in such a way that affirms individual choice while avoiding the problems noted here.

V. CONCLUSION—AFTER AUTONOMY?

I conclude with two brief contentions that may point to productive avenues for further consideration. First—and simply—in order for a doctor–patient relationship to be productive, the wills of two individuals must be aligned to some degree. “Hyper-paternalistic orientation” and “absolutized autonomy” approaches are actually similar in the sense that both tend toward totalizing the will of one party and subordinating the will of the other. All that liberalism has ended up asserting is a new form of illiberalism: substitution of tyranny of patient choice—something that is good neither for the non-expert patient nor for the caring doctor—for the supposed tyranny of the doctor. In situations of difficulty or conflict, what we *really* need to know, according to both of these accounts, is who is in power in the doctor–patient relationship. A richer account of ethics recognizes that the dynamic of mature human relationships is subtler than these accounts suggest. Here Richard Bauckham’s reflection on human freedom “in the crisis of modernity” is apropos:

The image of adolescence or “coming of age” . . . has been frequently used for the whole human project of the modern age . . . This was conceived as breaking free from dependence on nature through technological mastery of the natural world and as self-liberation from religious domination by assertion of independence from God . . . we could give [this account] the most generous . . . interpretation if we saw it in terms of an adolescent assertion of independence that has so far failed to mature into adult reappropriation of the relationships that have been repudiated. In appropriating freedom, modern humanity has not yet recognized that this very freedom is rooted in dependence . . . With the immaturity of an adolescent, modern humanity has absolutized its independence. Confusing belonging with dominion or ownership, it has failed to integrate the freedom it has asserted into new forms of belonging . . . As Nicholas Lash puts it, “It is surely time to learn the discipline of adulthood, the transcending of autonomy in community and finitude.” (2002, 178)

The image of adolescence is helpful here, I think. If we momentarily accept Beauchamp and Childress’ picture of pre-twentieth-century medicine

as inherently paternalistic—remembering that “paternalistic” most directly means “fatherly”—we might see the autonomy movement as a kind of adolescent step beyond vulnerable, childlike dependence. Also we might hope that adolescence will give way to adulthood: where conflicts of will need to be addressed via cultivation of reciprocal trust, healthy negotiation, and forbearing communication, instead of through recourse to appeals to power. The good doctor is neither a parent nor a government but neither is he or she a servile technician; the desires of the patient cannot be disregarded, but neither can they be seen as the ultimate criterion for determining proper action by the doctor. The doctor’s freedom to seek patient health is possible only through the patient’s initiation, enduring trust, and consent; reciprocally, the patient’s freedom to achieve proper medical care is possible only through the doctor’s free offering of expertise and genuine care. In other words, the space between doctor and patient is primarily one of clinical relationship, rather than political negotiation. In my view, this reorientation might suggest fruitful ways forward.

The evident appeal of Beauchamp and Childress’ principled approach to freedom in the doctor–patient relationship lies in its apparent logical straightforwardness. Such an approach is, however, doomed to work differently in practice than it does in theory because it fails to take into account the fundamental ambiguity that underlies mature relationships: that they must exist as a dialectic of two freedoms and not the subjugation of one will by another. What really matters, for good medicine and good autonomy, is not the question of who holds power but the question of how to foster what the British Medical Association has called “trust and reciprocity” between doctor and patient (English et al., 2004, 43). This understanding preliminarily suggests that a richer autonomy account cannot take the form of a simple, top-down, ethical principle. Most basically, a richer account of autonomy recognizes that autonomy is reciprocal and therefore that patient autonomy only exists *through* and *dependent on* the doctor’s free choice. To again quote Bauckham: “There is no human independence that is not rooted in a deeper dependence” (2002, 42).

Finally, I briefly return to where I began: the commonplace historical–philosophical sketch that purports to identify the source of the autonomy–paternalism dichotomy of the twentieth and twenty-first centuries. Critical analysis of this narrative is a vital precursor to richer dialogs that move beyond the impoverished debates within which we currently find ourselves entrenched. Is Beauchamp and Childress’ apparent assumption that the problem of paternalism was a disastrous flaw that somehow went unnoticed for 2500 years in an apparently continuous stream called “the Hippocratic tradition” convincing? Or might an alternative sketch be more helpful? What if the problem of paternalism is a corollary of Western society’s invitation to the doctor to fill the role of priest? If so, might it be that the ever-evolving autonomy principle is not merely a developing response to a properly understood problem,

but rather a notion destined for self-contradiction because it misidentifies the root of the problem of paternalism: medicine's bloated—albeit well-meaning—attempts at “biopsychosocial-spiritual” care?

NOTES

1. Jennings' (2009) “Autonomy,” cited below, is a very helpful illustration of this general line of argument. Secker's (1999) article, also cited, shows clearly the disparity between Kant's own thought and how he is often used in the bioethics conversation.

2. It is obvious, but still imperative to note, that these conclusions resemble and are deeply informed by Alasdair MacIntyre's analyses of post-Enlightenment moral language as explicated in *After Virtue*.

REFERENCES

- Bauckham, R. J. 2002. *God and the Crisis of Freedom: Biblical and Contemporary Perspectives*. Louisville, KY: Westminster John Knox Press.
- Beauchamp, T. L. 2004. Does ethical theory have a future in bioethics. *Journal of Law, Medicine and Ethics* 32(2):209–17.
- Beauchamp, T. L., and J. F. Childress. 1979. *Principles of Biomedical Ethics*. 1st ed. Oxford, United Kingdom: Oxford University Press.
- . 1994. *Principles of Biomedical Ethics*. 4th ed. Oxford, United Kingdom: Oxford University Press.
- . 2013. *Principles of Biomedical Ethics*. 7th ed. Oxford, United Kingdom: Oxford University Press.
- Berlin, I. 1969. *Four Essays on Liberty*. Oxford, United Kingdom: Oxford University Press.
- Callahan, D. 2000. *The Troubled Dream of Life: In Search of a Peaceful Death*. Washington, DC: Georgetown University Press.
- Conly, S. 2012. *Against Autonomy: Justifying Coercive Paternalism*. Cambridge, United Kingdom: Cambridge University Press.
- English, V., G. Romano-Critchley, J. Sheather, and A. Sommerville. 2004. *Medical Ethics Today: The BMA's Handbook of Ethics and Law*. 2nd ed. London, United Kingdom: BMJ Publishing Group.
- Hope, T., J. Savulescu, and J. Hendrick. 2008. *Medical Ethics and Law: The Core Curriculum*. 2nd ed. Edinburgh, United Kingdom: Churchill Livingstone Elsevier.
- Jennings, B. 2009. Autonomy. In: *The Oxford Handbook of Bioethics*, ed. B. Steinbok, 72–89. Oxford, United Kingdom: Oxford University Press.
- Kant, I. 1993. *Grounding for the Metaphysics of Morals*. Trans. J. W. Ellington. 3rd ed. Indianapolis, IN: Hackett Publishing Company.
- . 2009. *Religion within the Bounds of Bare Reason*. Trans. W. S. Pluhar. Indianapolis, IN: Hackett Publishing Company.
- Keenan, J. F. 2012. The case for physician-assisted suicide? In *On Moral Medicine: Theological Perspectives in Medical Ethics*, eds. M. T. Lysaught and J. J. Kotva, 3rd ed., 1085–8. Grand Rapids, MI: Eerdmans.
- Lysaught, M. T. 2004. Respect: Or, how respect for persons became respect for autonomy. *Journal of Medicine and Philosophy* 29(6):665–80.

- Mappes, T. A., and D. DeGrazia 1996. *Biomedical Ethics*. 4th ed. New York: McGraw-Hill.
- McCormick, R. A. 2012. Physician-assisted suicide: Flight from compassion. In *On Moral Medicine: Theological Perspectives in Medical Ethics*, eds. M. T. Lysaught and J. J. Kotva, 3rd ed., 1082–4. Grand Rapids, MI: Eerdmans.
- Meilaender, G. 2013. *Bioethics: A Primer for Christians*. 3rd ed. Grand Rapids, MI: Eerdmans.
- Mill, J. S. 2005. *On Liberty*. New York: Cosimo Classics.
- Otswald, M. 1982. *Autonomia: Its Genesis and Early History* Chico, CA: Scholars Press.
- Pellegrino, E. D., and D. C. Thomasma. 1993. *The Virtues in Medical Practice*. New York: Oxford University Press.
- Pink, J., L. Jacobson, and M. Pritchard. 2007. The 21st century GP: Physician and priest? *The British Journal of General Practice* 57(543):840–2.
- Preamble to the Constitution of the World Health Organization as Adopted by the International Health Conference, New York, 19–22 June 1946; Signed on 22 July 1946 by the Representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and Entered into Force on 7 April 1948.
- Savulescu, J. 2006. Ethics: Conscientious objection in medicine. *British Medical Journal* 332(7536):294–7.
- . 2007. Autonomy, the good life, and controversial choices. In *The Blackwell Guide to Medical Ethics*, eds. R. Rhoder, L. P. Francis, and A. Silvers, 17–37. Malden, MA: Blackwell Publishing.
- Secker, B. 1999. The appearance of Kant's deontology in contemporary Kantianism: Concepts of patient autonomy in bioethics. *Journal of Medicine and Philosophy* 24(1):43–66.
- Sulmasy, D. P. 2002. A biopsychosocial-spiritual model for the care of patients at the end of life. *The Gerontologist* 42(1):24–33.
- Veatch, R. M. 1989. Medical ethics: Its implications for the development of professional ethics. In *Values in Teaching and Professional Ethics*, ed. C. T. Mitchell, 37–52. Macon, GA: Mercer University Press.
- Verhey, A. 2012. The Bible and bioethics: Some problems and a proposal. In *On Moral Medicine: Theological Perspectives in Medical Ethics*, eds. M. T. Lysaught and J. J. Kotva, 3rd ed., 97–113. Grand Rapids, MI: Eerdmans.
- Wolfe, A. 2001. *Moral Freedom: The Search for Virtue in a World of Choice*. New York: W. W. Norton.
- Wolff, R. P. 1986. *The Autonomy of Reason: A Commentary on Kant's Groundwork of the Metaphysic of Morals*. Gloucester, United Kingdom: Peter Smith.