Melissa Rees
Siena College

Patient Autonomy and Withholding Information*

penultimate draft, please cite published version

Abstract: Disclosure in clinical practice is aimed at promoting patient autonomy, usually culminating in patient *choice* (e.g. to consent to an operation or not, or between different medications). In medical ethics there is an implicit background assumption that knowing more about (X) automatically translates to greater, or more genuine, autonomy with respect to one's choices involving (X). I challenge this assumption by arguing that in rare cases withholding information can promote a patient's autonomy (understood as the capacity for rational choice in alignment with one's values and goals).

Keywords: Autonomy, Consent, Disclosure, Information, Understanding

I. Introduction

Medical practitioners exercise, to greater or lesser extent, *discretion* in what they disclose to their patients. I begin with a vignette from my own life.

A friend and I have the same disability, one that commits us both to taking medication for the rest of our lives. This means we, on occasion, discuss our medication schedules, compare our experiences, and use each other as resources. On one occasion, my friend told me about a new medication—the only medication that (to this day) has successfully brought him relief. I was, of course, thrilled for him, and considered asking my own doctor whether I ought to consider taking it as well.

I subsequently learned—from a tv commercial—that the medication my friend was on had a risk of a permanent negative side-effect. This turned me off from the medication entirely. However, in a moment of morbid curiosity intermixed with no small amount of anxiety, I looked up the potential negative side-effects for the medication I was already taking. I learned that it too came with a risk of that same negative side-effect.

Did my doctor wrong me by failing to disclose this information to me? I will argue that he did fail to discharge a duty of disclosure (and so, did wrong me) but that this failure did not render my consent to take the medication invalid. I will also argue that, in cases different from mine, it can be possible to *support* a patient's autonomy by withholding information. When patients explicitly and forcefully express a goal of self-management (of a certain kind) practitioners can support the patient's autonomy by withholding information. I conclude by arguing that supporting autonomy through nondisclosure is compatible with informed consent.

The ethics of disclosure is a broad category in biomedical ethics. It ranges from issues like the one above—disclosing to a patient the potential negative side-effects of a medication—to issues about disclosing information from genetic testing or biological features of a fetus. In

what follows I will begin with some of the contours of the ethics of disclosure: the moral aim of disclosure, and some of the conditions that have been argued are necessary for discharging one's obligation to disclose. Ultimately, I will argue that there is a specialized subset of cases in which patient autonomy is, perhaps surprisingly, *promoted* by withholding information.

II. Autonomy

One aim of disclosure in clinical settings is to promote patient autonomy, usually culminating in either the tokening of informed consent or withholding of consent to a medical intervention. (A token—i.e. instance—of consent in medical settings is paradigmatically expressed through verbal or written affirmation). The general idea is that the choice to take a medication, or undergo a surgery, can only be a genuine expression of one's autonomy if the medical practitioner has appropriately informed the patient about the nature of the choice they are making.

The Principle of Respect for Autonomy—popularized by the Beauchamp and Childress' *Principles of Biomedical Ethics*—states that promoting a patient's autonomy is a pro-tanto duty. That is, it is a duty that ought to be discharged, albeit one that needs to be weighed against other (sometimes competing) duties, e.g. the duty to promote the welfare of the patient, or the duty to avoid harm. However, autonomy is notoriously difficult to define, and is correspondingly difficult to identify (and so, respect) in practice. Dive and Newson note that despite its central role in bioethics there, "is no consensus as to how we should understand it beyond a general notion of self-determination" (Dive & Newson, 2018, p. 171). This lack of consensus might not be so worrisome when one observes that biomedical ethics covers a vast range of clinical applications, and what it means to respect someone's autonomy is likely to be sensitive to the specific context of clinical encounter. Perhaps promoting someone else's self-determination looks quite different when a patient is contemplating medical assistance in dying, as compared to another patient who is contemplating reconstructive surgery on their hip. If that is right, then it seems perfectly appropriate to begin with a somewhat flexible understanding of autonomy as a starting off point, before attending to the particulars of the clinical encounter.

Indeed, Beauchamp and Childress have suggested that, "appropriate criteria of substantial autonomy are best addressed in a particular context, rather than pinpointed through a general theory of a substantial amount" (Beauchamp & Childress, 1994, p. 124). Their account nevertheless sets a firm foundation, subject to further specification. An autonomous choice requires *intentionality* (the act must be deliberate and not be accidental), *understanding* (the subject must sufficiently comprehend information relevant to the decision being made), and freedom from *controlling influences* (whether that influence is exerted externally, as is the case in coercion, or internally, as is the case of mental illness or addiction) (Beauchamp & Childress, 2019, p. 102). The above picture is, however, not yet complete. Although it captures features of rational choosing—i.e. these three criteria are necessary for a patient to make medical decisions for their own reasons (good or bad)—there is another expression of self-governance that is elsewhere called *authenticity* (Brudney & Lantos, 2011). Patients also express self-governance when they "form [them]selves at the deepest level so that our very desires are things we can endorse and that lead us to choices that consistently reflect those desires" (Brudney & Lantos, 2011, p. 221). This process of 'finding' or constructing ourselves allows us to find our activities meaningful—a kind of self-governance that is important for humans in all stages of life to develop and maintain but is particularly visible in

early adolescence[1]. Autonomy, then, requires a capacity for self-governance that is expressed through authentic rational choices, ones that reflect the patient's values and goals. This understanding of autonomy—the capacity for *authentic* rational choosing—is crucial to clinical encounters. When a patient chooses in a way that runs counter to their long standing, explicitly expressed, values and goals, one might rightly worry that the choice is not fully autonomous *even if* the choice is intentional, informed, and free from coercion.

In "Medical Ethics Needs a New View of Autonomy," Rebecca L. Walker identifies core cases of compromised autonomy that are not well characterized as a violation of any of Beauchamp and Childress's three conditions—*intentionality*, *understanding*, and *freedom from controlling influences*. That is, Walker identifies cases in which a subject's choice or action meets each of Beauchamp and Childress's three conditions, but the subject's choice or action is non-autonomous nevertheless. In addition to worries about authenticity like the one already mentioned ("…actions and choices that are deeply at odds with the settled and reflective ("authentic" if you like) self"), Walker introduces another paradigmatic failure of autonomy: "failures of the will to motivate one to action" (Walker, 2008, p. 597). Elucidating this kind of failure requires the introduction of a few technical terms.

A first-order state (e.g. belief, desire, intention) is one which takes as its object some proposition or state-of-affairs in the world. For example, I might believe the proposition that "it will rain tomorrow," or I might desire to finish reading the book on my bedside table, or I might intend to bring about the state of the world such that my desk is well organized. A second-order state takes as its object one or other of those first-order states. For instance, I might *believe that* "I believe it will rain tomorrow" or *desire to* have the desire to finish the book.

Harry Frankfurt introduces the notion of a second-order *volition* as a distinctive kind of second-order state, one that is essential for being a person and so essential for autonomous action. A second-order *volition* is a second order desire wherein one wants a particular first order desire to provide the motive for one's action. if I want my desire to finish the book to be *effective,* to determine my action, then I have a second-order volition directed at my first order desire to finish the book. By contrast, if I merely desire that I desire to finish the book (perhaps I think it is in general good to want to finish what you've started) but do not want my desire to finish the book to move me to *act*, then I have a second-order desire but not a second-order volition. In Frankfurt's words, a second-order volition is such that, "[Person] A wants the desire to X to be the desire that moves him effectively to act… He wants this desire to be effective-that is, to provide the motive in what he actually does" (Frankfurt, 10). On this view, my action is autonomous so long as the desire that moves me to act is endorsed by a second-order volition—my "true" self—regardless of the content or origin of the desire guiding my action. This is a procedural account of autonomy, part of a family of views that claims an action is autonomous if it is the result of the right kind of process.

Gerald Dworkin also has a procedural account of autonomy: regardless of *what* I desire, so long as I have a capacity for reflection at the second-order and my identification with (i.e. endorsement of) one or other of my first order-states has not been unduly influenced (it is "procedurally independent" e.g. does not result from manipulation, deception, hypnosis, etc.)

[1] See (Mullin, 2014) for a discussion of the parental responsibility to aid children in cultivating autonomy, understood as "a capacity to guide their activities to conform with goals the children themselves find meaningful" (p. 414).

my choice is made autonomously. "[A]utonomy is conceived of as a second-order capacity of persons to reflect critically upon their first-order preferences, desires, wishes, and so forth and the capacity to accept or attempt to change these in light of higher-order preferences and values" (Dworkin, 1988, p. 20). Procedural accounts of autonomy do not put any constraints on the *content* of the desire or intention that guides action—so long as it is a product of right procedure, the action is autonomously made.

A failure of the will to motivate action can be characterized as procedural failure—one's volitions and actions are not in accordance. There is philosophical disagreement about the details of such cases but what is common to weakness of will cases Robert Audi states, is "the notion of action against one's better judgement. The idea underlying this notion is that to act against one's better judgement is to do something intentionally… while in some sense aware of one's judging that doing something else would be better" (Audi, 1990, p. 270). Although I do not share Frankfurt's metaphysical commitment that identifies second-order states as the "true self"[2] this framework helps clarify what is meant by "weakness of will," second-order endorsement, and disavowal.

Of course, not all accounts of autonomy are procedural. Walker, for example, endorses a more substantive account of autonomy—one that does not prescind away from the content of our values, desires, and intentions. She argues that the most plausible reason why we are morally obligated to respect autonomy (in bioethics, and generally) is because our capacity for autonomy "[allows] us to create (or sustain) moral value or valuation and to act and choose in accordance with that value or valuation" (604). It is not enough for the agent's decision to reflect alignment between their higher-order preferences and their first-order states, there are further constraints on *what* those desires are, for the agent to be judged as truly autonomous. It would be difficult to understand why we are morally obligated to respect or promote an agent's autonomy if that translated to an obligation to e.g. respect or promote an agent's reflectively endorsed and procedurally independent decision to participate in their own oppression. However, if there are content-related constraints on autonomous action (in the loosest terms, that autonomy requires a sensitivity to goodness and badness[3]), the general sentiment that we have a moral obligation to promote autonomy is not mysterious.

This author endorses a 'weak' substantive account of autonomy that does not specify content beyond the negative requirement that a subject does *not* have a suite of desires or volitions that globally deny the legitimacy of their first-order states as reasons for them to act. As an illustration, a person who has reflectively endorsed and procedurally independently decided that their spouse will make all their decisions for them (i.e. they deny their *own* first order states any motivational role in guiding action) will not count as autonomous. I follow Paul Benson in the observation that there is a connection between truly autonomous agents and their reasons for acting that "does not permit their autonomy to be determined by things entirely beyond the reach of their motivational capacities" (Benson, 1991, p. 402). An autonomous agent's reasons for acting must be in some sense *their own*, and giving one's own desires and intentions a role in guiding one's actions requires sufficient self-regard.

---

[2] See also pg 14-15 of Catriona Mackenzie and Natalie Stoljar's *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* for a summary of objections to "hierarchical" or "structural" accounts of autonomy (Mackenzie & Stoljar, 2000).

[3] See also (Stoljar, 2018) for discussions of "normative competency".

That is not to say that autonomous agents are essentially "individual, self-made, independent, self-interested, isolated and insulated, and so forth" (Childress & Quante, 2022, p. 142). As Childress observes in his discussion of relational autonomy, "Self-rule, self-governance, or self-determination does not mean self-creation… we need to recognize how supportive social conditions foster autonomous action" (Childress & Quante, 2022, p. 142). Interpersonal relations help us shape and identify our values, beliefs, desires, intentions, and so forth. A supportive social environment is a pre-condition for autonomy and recruiting the assistance of intimate friends and family can in some cases be vital for fostering the autonomy of a patient and maintaining trust between patients and practitioners (a topic that will be discussed at greater length in section VIII).

I believe that the key insight of this paper—that patient autonomy can in some cases be supported by withholding information—is compatible with either procedural, substantive, or weakly substantive accounts of autonomy. However, detailed illustration of that claim must be reserved for future work.

III. Consent

One highly visible expression of autonomy is, of course, informed consent. Informed consent is often the culmination of autonomous decision-making that reflects what the patient cares about. In a classical liberal framework consent has the 'moral magic'[4] of granting permissions, rendering what would otherwise be a violation into a permissible intervention. Specifically, it waives person B's right for person A to abstain from some action, whether that be entering one's house, or slicing into one's body.

Informed consent—although very important—does not by itself fully determine the limits of moral permissibility. I follow Onora O'Neill in her observation that the importance of consent-based language in medicine is to guard against misconduct (e.g. deception and coercion) more generally, rather than the "libertarian tendency in medical ethics [that] sees informed consent as *necessary* and *sufficient* justification for action" (O'Neill, 2003, p. 5). That is, the presence or absence of consent does not exhaust the moral landscape although it often can be a good guide for assessing the moral (im)permissibility of an intervention. On the one hand, it seems plausible that person B might waive their right for person A to refrain from needlessly amputating their leg and, nevertheless, person A wrongs person B if they perform the amputation. If that is right, consent is not sufficient for moral permissibility. Similarly, there may be cases in which person B does not token consent where we nevertheless judge person A's action upon B to be permissible. For example, it may be morally permissible for person A to resuscitate person B when B is unconscious and not capable of tokening consent. If that is right, consent is not necessary for moral permissibility.

Respecting someone's autonomy and securing valid consent are clearly related—the former supports the latter. They are nevertheless conceptually distinct. That is, one might give valid consent even in a context when one's autonomy is not being respected, and a consent-receivers respect for the consent-giver's autonomy does not guarantee the validity of the consent-giver's consent. For example, a patient might validly consent to take oral contraception even if, during the consultation, the medical practitioner or a family member did not respect the patient's autonomy (perhaps they exerted pressure on the patient to choose otherwise). A patient might also appear to consent to a surgery in an environment highly

[4] This phrase is owed to Heidi Hurd's legal analysis of consent (Hurd, 1996).

supportive of their autonomy—let us stipulate that their values were discussed and respected and that they were broadly and clearly informed. But perhaps despite being presented with evidence to the contrary, the patient retained a false belief that the procedure was reversible. This would vitiate their consent, even in an environment where their autonomy was being respected.

The purpose of this article is to illustrate that withholding information can *support* autonomy. In light of the distinction between respecting autonomy and securing valid consent, however, one might *agree* that in a given clinical encounter autonomy was supported by withholding information and *still worry* that consent in that case was not valid. I address this worry in section VIII.

In the following section I will scrutinize the plausible claim that consent is valid only if all *difference making* information has been disclosed. That is, disclosing difference making information is a necessary condition for valid consent: call this the *difference-making criterion*. This captures the intuitive idea that certain pieces of information must be disclosed precisely because (i) with the presence of that information the patient is more likely to make one choice, and (ii) had the information been absent, the patient is more likely to have made an alternative choice. I will first argue that some kind of difference-making criterion for disclosure can clarify the peculiarity of my opening vignette, and argue for the claim that withholding information can support autonomy *even when* difference making information is withheld.

IV. Disclosure vs. Understanding

A duty to disclose is not *ipso facto* a duty to guarantee that a patient has reached a state of knowledge or understanding. That is, one might fulfill a duty to disclose information but nevertheless fall short of guaranteeing that the patient fully understands that information. Given that it is not possible to disclose *every* piece of relevant information in clinical practice nor is an absolute level of understanding attainable for even the most epistemically skilled patients, I will refer to 'broad disclosure' and 'broad understanding' rather than more absolute terms like 'full' disclosure or 'complete' understanding. In his paper "Informed Consent, Disclosure and Understanding," Tom Dougherty helpfully distinguishes between two *informational conditions* for valid consent: the Facilitative Disclosure Condition, and the Context Understanding Condition (Dougherty, 2020).

Dougherty argues that consent-receivers have "facilitative duties" and "[a]s you might expect, the consent-receiver's facilitative duties include a duty to disclose relevant information" (Dougherty, 2020, p. 125). For Dougherty, facilitative duties are aiming at *relative improvements* in a person's "position" for giving consent. I am here assuming that positionality is a spatial metaphor for the consent-giver's epistemic states, practical skills, and perhaps (although this is not mentioned) affective or emotional states. The facilitative duty to disclose can be discharged even if the consent-giver's position is improved but nevertheless falls short of an ideal position. The facilitative duty to disclose can be discharged even if the patient does not attain a broad understanding of the risks and benefits of a therapy. Perhaps this means that some, but not all, information about the therapy was disclosed. Or, this might mean that all information was disclosed, but the patient did not understand all of it. It might also mean that a subset of the total informational body the consent-receiver had access to was disclosed, and further, the patient understood only a subset of *that*.

In the context of partial understanding, for the facilitative duty to disclose to be discharged, the consent-giver must be able to "take responsibility for their ignorance of the nature of the risks" (Dougherty, 2020, p. 131). The idea is that if a patient fails to advertise their lack of understanding, or, "it becomes clear to the consent-giver that they cannot grasp certain information, and yet they still wish to persist with a procedure," then the consent-receiver has met the facilitative disclosure condition (Dougherty, 2020, p. 130).

However, more needs to be said about what it is to be in a sufficiently good position for giving consent. Beyond the attempt to enact the relative improvements that facilitative duties require, one might also think—rightly—that there is a need for *some* absolute threshold that must be met for consent to be validly given. Dougherty argues that this need is captured by the Context Understanding Condition, which states that, "valid consent to an action also requires understanding this action's purpose, risks, and benefits, and similar information about any alternative options" (Dougherty, 2020, p. 132). This condition covers the "major risks and benefits, along with alternative options" of an intervention, but does not include "everything that the consent-giver would find helpful in making their decision" (Dougherty, 2020, p. 143).

This framework can shed light on the moral significance of my opening vignette. The intuition that I did not validly consent to take my medication can be explained by what appears to be a violation of the context understanding condition: I did not understand one of the risks associated with my medication. Nor was I put in a position to assume responsibility for this ignorance—thus my medical practitioner also violated the facilitative duty to disclose. That is, one might reasonably infer from the lack of respect for my autonomy (i.e. the failure to meet the duties of disclosure) that my consent was vitiated.

However, I think that something like a difference-making criterion can also illustrate why one might equally think my consent was valid. As it turns out, *all* my available options were on a par with regards to the negative side-effect. Every treatment option available to me came with the risk of this one negative side-effect. Further, my commitment to taking *some* medication for my condition was fixed (the option to take *no* medication was not, by my lights, an option at all) and my doctor was aware of that. Consequently, disclosing the risk of the negative side-effect was predictably not difference-making with regards to my choice.

Minimally, this is simply another illustration of the claim that consent can be valid even in contexts where one's autonomy has not been respected. The lesson I draw from this vignette pertains to the status of the difference-making criterion. The subset of information that makes a difference to us is determined entirely by our commitments, desires, and values.

Sometimes information is difference-making because it *threatens* or *undermines* our goals and commitments. I will go on to suggest that in such cases, it is not a necessary condition for consent that all difference-making information be disclosed. Under special circumstances, withholding information (*even if* it would have made a difference in patient choice) does not always undermine autonomy nor result in the vitiation of consent.

V. Information and Autonomy

There is an implicit background assumption in literature on the ethics of disclosure that meeting a higher standard of disclosure is always supportive of the patient's self-governance.

However, I reject the claim that knowing more about (X) automatically translates to greater, or more genuine, autonomy with regards to (X). I challenge the claim that meeting higher standards of disclosure always promotes patient autonomy. It is an argument to the effect that disclosure interacts with autonomy in very complicated ways, and if that is right, then it is at least conceptually possible to identify circumstances in which opting for higher standards of disclosure might itself exert unwelcome influence on a patient's practical attitudes toward an intervention.

Consider the following case:

> **Weight Gain**: patient (P) is seeking treatment for a condition that moderately interferes with their ability to pursue activities they find meaningful. The patient is also very weight-conscious, and if they learn a medication has a risk of weight gain, they will not take the medication. However, (P) does not endorse their aversion to weight gain.

The patient's weight-consciousness might retain its unendorsed status for any number of reasons. Perhaps they recognize in themselves weight-related stigma, perhaps they believe that their weight-consciousness originates from an irrational fear or anxiety, or perhaps they know that weight gain would not impact their health negatively and value their health more than their body size. Whatever (P)'s reasons are, the first order desire not to gain weight is unendorsed but nevertheless motivationally powerful. (P) is aware that their desire to avoid weight gain will move them to act—in this case, to refuse a beneficial treatment.

Even limiting the discussion to *one* side-effect (weight gain) we can see multiple clinical applications. In cases where rheumatoid arthritis is interfering with a patient's ability to garden or chop vegetables, providing broad disclosure regarding the various options involving corticosteroids (P) might forego such treatments if their unendorsed desire to avoid weight gain is motivationally strong enough. The stakes can, however, be raised. Perhaps (P) has bipolar disorder, but their unendorsed aversion to weight gain would lead them to choose against a preventative use of anti-psychotic treatment. This example is particularly pertinent given the high co-morbidity between bipolar disorder and disordered relationships to food and body image.

The clinical applications for this understanding of disclosure and autonomy expands considerably when one considers the many other side-effects amenable to this analysis and the clinical condition for which they are relevant (I discuss this issue of relevance—the extent and limits of this analysis—in section VII).

I argue that under such conditions, the medical practitioner counseling patient (P) can ethically withhold the information that a proposed treatment option has a risk of weight-gain.

One might worry that something has gone wrong here. After all, if the patient *had known* the potential side-effects, they would not have wanted to take (X). And if that is true, isn't it clear that withholding information in this case necessarily undermines the patient's autonomy? I would like to suggest that it does not. Rather, the moral permissibility of this kind of case depends on a distinction Pamela Hieronymi makes between evaluative and managerial control.

VI. Evaluative and Managerial Control

Hieronymi argues that there are two kinds of agency we can have with respect to our beliefs and intentions. We exercise evaluative control over a belief or an intention when we *form* or *revise* our answers to the question whether to believe p or intend to ϕ. For Hieronymi, when we have settled the question whether to believe p or intend ϕ, and we are "committed" to believing p or intending ϕ (i.e. when we are satisfied by the reasons we take to bear positively on our answers to these questions) we have "embodied" our answers to those questions. When an attitude embodies an answer to the question (whether to believe, whether to intend) positively, we therein believe or intend.

This is not an ordinary account of control or agency. "The forming and revising of beliefs and intentions is not voluntary nor does it require the same kind of reflective distance or awareness" (Hieronymi, 2009). When we exercise evaluative control, our beliefs and intentions do not stand at a reflective distance from us in our thought and we cannot form or revise them for any reason whatsoever but, "only for reasons that we take to show the belief true or to bear on whether to perform the action intended" (Hieronymi, 2009). Nevertheless, in settling the answer to the question whether to believe p or intend ϕ we do exhibit a kind of control over our minds. We *change* our minds—albeit not voluntarily nor at a distance—and that is a form of control.

Hieronymi takes it that forming and revising our beliefs—settling the question whether to believe p or intend to ϕ either positively or negatively—is the "ordinary and most fundamental way" in which we exercise control over our attitudes of belief and intention. It is possible to give an example: imagine a patient who has the open question whether or not to undergo surgery. There are many things that they will take to bear upon their answer to that question. When those things are presented—their diagnostic test results, the experience in their body, information about the risks and benefits of the operation and the alternatives relative to their values and aspirations—Hieronymi thinks that they exercise control (in this case, over their intention whether or not to undergo surgery) when they settle the question either positively or negatively. Although this is the fundamental way in which people exercise control over their beliefs and intentions—it is common practice for us to form or revise our epistemic and practical attitude in response to what we take to be reasons that bear upon believing or disbelieving, intending or not intending—this is not a widely recognized form of 'control' both because it is involuntary and also because we do not reflect upon the beliefs and intentions themselves. When we come to believe that it will rain today, we do not reflect upon our *belief,* we reflect upon features of the world that bear upon the question whether it will rain. When we come to intend to undergo surgery, we do not reflect upon our intention, we reflect upon the features of the world that bear upon our decision (our values, the diagnostic tests, information about the surgery). That is, we do not stand at a distance from our belief when we form a belief—rather, we *embody* the belief. And we do not stand at a distance from our intention when we form a plan of action—we *embody* that intention.

By contrast, we can exhibit managerial or manipulative control over our beliefs and intentions, canonically, "by taking steps that you can predict will provide you with convincing reasons for the answer embodied in the attitude" (Hieronymi, 2009). Hieronymi uses the analogy that managerial control is the same kind of control you might bear to ordinary objects around you. I exert managerial control over my coffee mug by placing it firmly in the center of the desk rather than placing it precariously near the edge, because I predict that the cup will interact with the environment in predictable ways (the edge of the

desk is precarious because it will, predictably, fall off). Exercising managerial control over a belief or intention involves acting upon it in ways that set up the belief or intention in predictable relations to the environment, in pursuit of our goals. If you have the goal of going to the gym, you might exercise managerial control over the intention to go to the gym by setting out your clothes the night before, making a bet with your partner that you will go, leaving yourself motivational notes around the house because you predict that the intention will relate to these features of the environment in predictable ways that are in support of your goal. We also exercise managerial control when we direct our attention in certain ways—perhaps by seeking out or avoiding new evidence. Importantly, this is the same kind of control that others can exert on us: it is the kind of control exerted when an agent is "taking actions designed to affect [our beliefs and intentions] according to our purposes" (Hieronymi, 2009). When another person texts you on the phone to assure you that they have arrived home safely, they are exerting managerial control over your belief in their safety. When a medical practitioner chooses to disclose or not disclose information to a patient, they are exerting managerial control over the patient's epistemic and practical attitudes. The ethical question is: when does an act of disclosure *interfere* with a patient's autonomy? The difference—which I hope is clearer now that Hieronymi's distinction between evaluative and managerial control has been introduced—has to do with the disclosure strategy's alignment or misalignment with *the patient's own* managerial goals.

Discussions of autonomy tend to focus primarily on patients' *evaluative control*. By opting for higher standards of disclosure one is aiding, promoting, supporting the patient's ability to exercise evaluative control, their ability to settle the question whether to take the medication or not. In this sense, there is a close relationship between the disclosure of information and the promotion of autonomy and for that reason the focus on patients' evaluative control is justified but results in an incomplete picture of how disclosure can interact with autonomy. A patient might have *managerial* goals that run at cross purposes to the well-intended medical practitioners' intention to aid, promote, or support the patient's evaluative capacities by maximally informing them.

Some patients might not want to know the side-effects of a medication precisely because that will 'engage' their evaluative control and settle the question "should I take this medication?" negatively. Insofar as they have the intention to take the medication, and do not want that intention interfered with, it can be ethical for a medical practitioner to withhold information about negative side-effects. This is true *even though* the patient would *not* take the medication (they would not settle the question whether to take the medication positively) if they knew the side-effects. By withholding information, the medical practitioner is promoting the patient's managerial control over their own intentions, even if the medical practitioner is circumventing the patient's evaluative control over their own intentions.

I suggest that because this relation between withholding information and promoting patient autonomy has not been theorized, we do not know how common it is for patients to possess managerial goals that are supported by withholding information. As such, this is an opportunity for discovery and for conversation. Of course, in all cases where patients have the managerial project of seeking out information failing to disclose would circumvent both the patient's evaluative control and their managerial control. Such a practice would be ethically impermissible.

However, the framework that casts questions about the ethical permissibility of different disclosure strategies *purely* in terms of promoting evaluative control—with no attention to different ways that autonomy can be promoted or stymied—is not sufficient.

VII. The Limits of Promoting Managerial Autonomy

The picture outlined above—wherein I have argued that in exceptional cases patient autonomy can be promoted by withholding information—has strong limits. I will address two pressing issues: limitations in how much a practitioner can know about their patient's managerial goals, and paradigmatic cases in which a practitioner should not promote the patient's managerial goals (even if they have succeeded in identifying them).

As already mentioned, patients may have the managerial goal to *seek information* that will help engage their evaluative control and settle the question: "should I take this medication?" or "should I undergo this surgery?". The practice of providing more extensive disclosure will of course be necessary in all such cases. So too is it appropriate to assume patients have the managerial goal of seeking information unless given strong, decisive, and explicit testimony to the contrary. There are also strong limits as to what kind of information is permissible to withhold—I suggest it will include things like risk of weight gain, dry mouth, itchiness and pain, and not side-effects like heart attack, stroke, or permanent paralysis. The boundary is a vague one and precisifying the boundary (to the extent that it is possible) is the task for future work. The relevant question here, is whether and how one could possibly identify the rare cases in which a patient has a managerial goal to remain uninformed with regards to some risk or set of risks associated with the intervention.

It is obvious enough that, in cases when the patients themselves are not aware of their own managerial goal to avoid certain information—perhaps it merely tacitly guides them, rather than being explicitly represented to them in their reasoning—medical practitioners cannot act in accordance with those goals, not should they attempt to. But what of the cases in which a patient has a clear interest (clear to themselves) to seek out some subset of information and avoid a different subset?

Medical practitioners cannot simply ask. This is nicely illustrated in Wertz and Fletcher's discussion of privacy and disclosure practices in medical genetics. Some patients do not wish to know if they have genetic risk, but geneticists cannot ask them directly with any specificity:

> Should the geneticist seek them out and tell them? What if they do not wish to know? There is no way for them to exercise the choice of not knowing, because in the very process of asking, "Do you want to know whether you are at risk for HD [Huntington's Disease]?" the geneticist has already made the essence of the information known (121).

If a medical practitioner asks, "do you want to know if this medication puts you at risk of X?" one can reliably infer that there is a risk of X from the content of the question. Call this *risk implicature*—wherein a proposition (e.g. that there is a risk of X) is implicated by the asking of a question.

The "right not to know" became a topic for direct interrogation as genomic information became more popularized in medicine in the 1990s (following the launch of the Human Genome Project) (Chadwick, Levitt, & Shickle, 2014). Topics of interest included whether individuals have a right not to know incidental findings of genetic risks after testing, and whether family members have a right not to know their genetic risks when a family member sought testing.

The "right not to know" of course has both defenders[5] and challengers[6]. Notably, Lisa Dive has argued that the "right not to know" is better conceived in terms of patient *preference* rather than patient *rights* in part because rights are often taken to override all other moral considerations, whereas preferences are one amongst many considerations in a clinical encounter[7]. Whether "preference" identifies the kind of committed self-management I appeal to in this article depends upon how the notion is deployed in context, but I am in strong agreement with the principle inherent in Dive's treatment of genomic ignorance: in her words, that "more information does not necessarily enhance autonomy" (Dive, 2021, p. 615).

How can medical practitioners come to learn what it is that their patient does not want to know, without introducing risk implicature? One might think that medical practitioners can learn a lot about their patient's managerial goals by assessing their patient's general attitudes toward risk itself. One background assumption that I am adopting here is that it is possible to consent to *gambles*. But, in order to consent to a gamble, one must have a sense of the odds.

Perhaps it is possible, then, to ask patients in a content-neutral way what their threshold for risk is. Perhaps some patients are happy in general to remain unaware of risks that affect one in ten thousand people, but not one in one hundred. Others may want to be informed of all risks even if likelihood is astronomical: one in one hundred million. Our risk aversion is idiosyncratic, and useful background information for a medical practitioner to know.

However, an entirely content-neutral assessment of risk-aversion is not sufficient. One might be happy to take the gamble that I will experience mild itchiness when it affects one in one hundred people, and am not happy to take a gamble that I will experience severe joint pain when it affects one in one hundred people. So content-specific risk aversion will also be idiosyncratic. Risk of weight gain will matter quite a lot to some, but not all. So too with drowsiness, or joint pain—likely to be determined by the patient's values, goals, their job, and general lifestyle.

Although knowing a patient's general attitude toward risk, and the patient's values, might be a good start toward knowing a patient's managerial goals, it still does not settle the question: does the patient wish to be informed?

Again, there will be some outcomes that are so severe that they must be disclosed regardless of a patient's managerial goals—death, for instance. But even for non-life-threatening

[5] See (Takala, 1999), (Andorno, 2004).
[6] See (Ost, 1984),(Rhodes, 1998).
[7] This point can also be found in Harris and Keywood's "Ignorance, Information and Autonomy" where they argue that even within a rights-based framework, any "right" to ignorance would be rivalled by rights to free speech and a right to decline the responsibility to make decisions for others when they are imposed on you (e.g. a patient who exercises their "right" not to know may thereby impose on their practitioner to make decisions for them, thus infringing their practitioner's rights) (Harris & Keywood, 2001). See also (McDougall, 2004) for a cautioning against rights-based frameworks for genetic ignorance.

outcomes, a risk averse patient who cares deeply about that outcome might nevertheless have a surprising managerial goal. For example, a highly risk-averse patient might care deeply about their weight, and nevertheless wish to exercise their managerial control by being uninformed about the risk for weight gain precisely because that is the kind of information that will decisively affect their evaluative control.

The best avenue for getting a sense of a patient's managerial goals, then, requires communication antecedent to any discussion of a particular intervention. Risk implicature occurs after an intervention (or interventions) have been introduced into conversation. If interventions Y and Z are already under discussion and a medical practitioner asks, "do you want to be informed about the risk of weight gain?" the patient will justifiably infer that Y and Z come with a risk of weight gain. Compare that with the kind of conversation that might take place in the context of an enduring diachronic relationship between the medical practitioner and the patient where no particular intervention is under discussion. The medical practitioner asks: "do you want *in general* to be informed of the risk of weight gain in our future meetings?" and "how high a risk would make weight-gain a difference-maker with regards to your decisions?" Only in these contexts—where patients express an explicit (and forceful) preference about disclosure practices antecedent to a discussion about a particular intervention—can a medical practitioner get a sense of a patient's managerial goals while also avoiding risk implicature. Of course, managerial goals can change over time. These kinds of "check-ins" would need to be ongoing. In situations where a medical practitioner does not have the time or resources to engage in such conversations, one must opt for broad disclosure. Another helpful resource for clinicians would be the community of people, in particular, close trusted friends or family that the patient surrounds themselves with. If a trusted friend is intimately aware of the patient's values, goals, and threshold for risk, the patient might nominate them as a person the clinician can turn to, to get a better sense of the patient's managerial goals without inviting the possibility of risk implicature.

Nevertheless, there will still be cases where a patient's managerial goals—even when they are properly identified—should not be taken up by the medical practitioner. Consider anti-vaccination populations who have decided, in advance, to avoid information about viruses (their incidence rates, and risks of long-term effects) in order to support their overall conviction to remain unvaccinated[8]. Perhaps such a patient expresses this preference forcefully and in an ongoing way. In this case, and in any case wherein the medical practitioner has reason to believe that the patient has *false beliefs* with regard to the risks or benefits of an intervention, the medical practitioner should not promote the patient's managerial control. To do so would amount to a *deceptive* disclosure practice because the medical practitioner (i) predicts that the patient has false beliefs, and (ii) withholding information leaves the false belief intact. This example is perhaps overdetermined because vaccination is a public health issue as much as it is a personal health issue. But other examples can be provided: whenever a patient has a managerial goal to avoid information about X, and a medical practitioner believes that this goal is partially motivated by a medical *misunderstanding*, the medical practitioner should not promote the patient's managerial goal to remain uninformed.

[8] Of course, the individual's commitment might be politically motivated, and receiving such information would make no difference to the patient's evaluative control. I am stipulating that in this case the individual's commitment to avoid information is motivated by their prediction that *were* they to know more they would change their mind.

VIII. Trust and Consent

It is now possible to return to the question: does supporting autonomy by withholding information undermine *consent*. On the view I have suggested, aiding a patient in pursuing actions more aligned with their higher-order endorsements and managerial goals by withholding information is supportive of their autonomy. However, this leaves open the question whether decisions that result from *these kinds* of expressions of autonomy are also expressions of valid consent. To recall Dougherty's formulation: "valid consent to an action also requires understanding this action's purpose, risks, and benefits, and similar information about any alternative options" (Dougherty, 2020, p. 132). This condition covers the "major risks and benefits, along with alternative options" (Dougherty, 2020, p. 143). My suggestion is that recognizing and respecting the patient's managerial goals (within the limits stated above) in fact gives medical practitioner's guidance as to how to interpret what counts as a major risk or benefit and what counts as a genuine alternative option. For a weak-willed patient with an unendorsed desire to avoid weight gain, withholding information in fact expands the space of alternative options according to what *they* reflectively judge to be a "major" risk or benefit. In this case, withholding information expands their options to include interventions that come with a risk of weight gain. The space of opportunity is furthermore not being dictated by a third party. It is a result of the patient's *self* management, made possible by the help and support of their practitioner. The *informed* part of informed consent is—I take it—supposed to (1) prevent coercion and manipulation, and (2) facilitate patient choice. On the former point, the patient is not being coerced or manipulated by their medical practitioner. On the latter point, patient choice is being facilitated by increasing their pool of options according to their own endorsed values.

There is a final worry. Will withholding information ultimately undermine the patient's trust in their medical practitioner, or in the medical establishment more generally?[9] This is particularly pressing given the wealth of information a patient might access from other sources. If a patient learns of a side-effect that was not disclosed (albeit at their own request) might the patient experience resentment or mistrust?

One insight from Childress's discussion of relational autonomy is the important (and perhaps underutilized) resource that intimate family and friends might play in their clinical experience. The possibility of mistrust might be mitigated in the following way: a clinician might write the withheld information down in a sealed packet, and present that packet to a patient's appointed intimate relation. Were the patient to come to learn information from other sources that had been withheld by their practitioner (and feel negative attendant emotions like mistrust, resentment etc.), they could then access the packet and see that the information was available there: information which clearly *would have been disclosed* had it not been for their own explicit instruction.

IX. Conclusion

Opening discussion on the claim that more information does not immediately translate into greater or more robust expressions of autonomy does not license making assumptions about patient's managerial goals—quite the opposite. The spirit of this argument is to open a dialogue about the ways in which medical practitioners can better come to know their patient's unique managerial goals, and so, promote their autonomy in perhaps surprising

[9] I thank an anonymous reviewer for raising this critique.

ways. Assuming that a patient wishes to remain uninformed is a more grievous ethical error than assuming that a patient wishes to be broadly informed. So, although I do take this to be a right account of information as it relates to autonomy, implementation must necessarily be cautious.

Works Cited

Andorno, R. (2004). The right not to know: an autonomy based approach. *Journal of Medical Ethics, 30*(5), 435-439.

Audi, R. (1990). Weakness of will and rational action. *Australasian Journal of Philosophy, 68*(3), 270-281.

Beauchamp, T. L., & Childress, J. F. (1994). *Principles of biomedical ethics* (4th ed.). New York: Oxford University Press, USA.

Beauchamp, T. L., & Childress, J. F. (2019). *Principles of Biomedical Ethics* (8th ed.). New York: Oxford University Press.

Benson, P. (1991). Autonomy and oppressive socialization. *Social theory and practice, 17*(3), 385-408.

Brudney, D., & Lantos, J. (2011). Agency and authenticity: which value grounds patient choice? *Theoretical medicine and bioethics, 32*(4), 217-227.

Chadwick, R., Levitt, M., & Shickle, D. (2014). *The right to know and the right not to know: genetic privacy and responsibility*: Cambridge University Press.

Childress, J. F., & Quante, M. (2022). *Thick (concepts of) autonomy: Personal autonomy in ethics and bioethics* (Vol. 146): Springer.

Dive, L. (2021). *From a right to a preference: rethinking the right to genomic ignorance*. Paper presented at the The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine.

Dive, L., & Newson, A. J. (2018). Reconceptualizing autonomy for bioethics. *Kennedy Institute of Ethics Journal, 28*(2), 171-203.

Dougherty, T. (2020). Informed consent, disclosure, and understanding. *Philosophy & Public Affairs, 48*(2), 119-150.

Dworkin, G. (1988). *The theory and practice of autonomy*: Cambridge University Press.

Harris, J., & Keywood, K. (2001). Ignorance, information and autonomy. *Theoretical medicine and bioethics, 22*(5), 415-436.

Hieronymi, P. (2009). Two kinds of agency. *Mental actions*, 138-162.

Hurd, H. M. (1996). The moral magic of consent. *Legal Theory, 2*(2), 121-146.

Mackenzie, C., & Stoljar, N. (2000). *Relational autonomy: Feminist perspectives on autonomy, agency, and the social self*: Oxford University Press.

McDougall, R. (2004). Rethinking the 'right not to know'. *Monash Bioethics Review, 23*(1), 22-36.

Mullin, A. (2014). Children, paternalism and the development of autonomy. *Ethical Theory and Moral Practice, 17*(3), 413-426.

O'Neill, O. (2003). Some limits of informed consent. *Journal of Medical Ethics, 29*(1), 4-7.

Ost, D. E. (1984). *The 'right'not to know*. Paper presented at the The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine.

Rhodes, R. (1998). *Genetic links, family ties, and social bonds: rights and responsibilities in the face of genetic knowledge*. Paper presented at the The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine.

Stoljar, N. (2018). "Feminist Perspectives on Autonomy". In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Online.

Takala, T. (1999). The right to genetic ignorance confirmed. *Bioethics, 13*(3-4), 288-293.

Walker, R. L. (2008). *Medical ethics needs a new view of autonomy*. Paper presented at the The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine.