

## Autonomy, consent and the “non-ideal” case<sup>1</sup>

Hallvard Lillehammer

### 1. Introduction

In much of the recent literature on the subject, autonomy is interpreted as having the capacity and freedom to be the primary judge and executor of how one’s life goes (see e.g. Dworkin 1998; Dworkin 1994; Beauchamp and Childress 2008; Korsgaard 2009; Radoilska 2013). In the case of the normal and competent human adult - sometimes identified with the enfranchised citizen of a modern democratic state - our capacity for self-governance can be thought of as grounding a constraint on what other people (including the state and its representatives) can legitimately do to us, thereby providing a rationale for consent requirements of various sorts (see e.g. Estlund 2007). On this view, the failure to elicit my consent in the context of some specific interaction is to fail to respect me as the autonomous, and thereby normatively qualified, agent I am. One obvious limitation of this explanation is that the practice of constraining behaviour by eliciting consent extends far beyond the domain of agents who satisfy the standard requirements of autonomous, self governing, rational agency (see e.g. McMahan 1996; Beauchamp 2005; Kittay 2005; Levy 2006; Nussbaum 2009; Lillehammer 2012). Cases in this category include persons with mental health

---

<sup>1</sup> Parts of this paper have been previously presented at a workshop on *The Theory and Practice of Informed Consent* at Oslo and Akershus University College (now Oslo Metropolitan University) in June 2017. I am grateful to the audience on that occasion for their questions; to Bjørn Hoffman for responding to a draft version; and to Edmund Henden for the invitation to participate. I am also grateful to Michael Garnett, David Owens and two anonymous referees for their comments on a subsequent draft.

difficulties involving self-destructive behaviour (such as substance addiction) or denial (such as anorexia); limited and diminishing capacity or competence (such as dementia)<sup>2</sup>; and a large variety of less diagnostically circumscribed behaviours, such as temporal myopia; inconstancy of preference; indecisiveness; ill temper; extensive confusion, anxiety, or severe depression. What explains the constraints on our (here presumed to be beneficent) behaviour towards such persons, and on what grounds could we be said to owe them a duty to elicit consent?

## **2. Consent in context**

Three preliminary points by way of clarification. First, in what follows I will not do justice to a number of the most contested aspects of the ethics of consent. For example, I will have little to say about the distinction between actual and hypothetical consent. I will generally assume that the consent we are interested in is actual, not merely hypothetical or imagined. Thus, if my students were to raid my bookshelf while I am away at a conference the fact that I might have agreed to lend them the books had they asked me to does not legitimate the raid. In making this assumption, I do not mean to imply that hypothetical consent never has justificatory force. On the contrary, I think it does (e.g. in some cases of forgetfulness or distraction.) Moreover, I will have little to say about the distinction between acts of consent

---

<sup>2</sup> “Capacity”, or “competence”, is a legally recognised status in many jurisdictions, and in the UK is widely understood as follows: an agent is considered to have capacity/be competent with respect to a decision if she i) can understand ‘the information’ relevant to making that decision; ii) can retain that information to make a decision; iii) can weigh up that information to make a decision; and iv) can communicate their decision based on weighing that information (See e.g. Alzheimer’s Society (GB), 2018). Why we should respect mere capacity or competence, so understood, is one of the questions addressed in this paper, although the domain of cases at issue can also be extended beyond the domain of capacity or competence as defined above (and in comparable documents).

and consent that is otherwise inferred by the fact that someone “goes along” with something, and is thereby regarded as having “tacitly” or “implicitly” consented to what could then be described as a “consensual act”. I will generally assume that the consent at issue involves some kind of act (whether linguistically articulated or otherwise). Thus, if my students were to raid my bookshelf while I am busy putting out a fire elsewhere in the building, the fact that for the moment I let them get on with it does not legitimate the raid either. In making this assumption, I do not mean to imply that interpersonal interactions without prior acts of consent are thereby illegitimate. On the contrary, I think the insistence that only acts of consent can legitimate intervention is frequently misguided (e.g. in cases where a gentle nudge will do); sometimes perverse (e.g. in cases where it would undermine fragile trust); and sometimes incoherent (e.g. if a client insists that every possible eventuality must be explicitly covered in a contract). Nor will I have anything to say about the fact that some legitimate consent is given by a “surrogate” (as when I authorise a professional to represent me as my “agent”, or when adults are placed in the position of making choices about the medical treatment of infants). Finally, I will have little to say about the fact that consent is referentially opaque. I will mostly assume that the parties in question could in principle agree about at least some of the basic and ethically relevant aspects of the question at issue (although there will be cases involving incapacity where this is not the case). Thus, if I consent to one of my students relieving me of “some” of my books, the fact that “some” is consistent with “all” in First Order Predicate Logic does not legitimate her making off with the lot. In making this assumption, I do not mean to imply that interpreting someone’s consent as extending further than the range of descriptions they have consciously entertained is always illegitimate. On the contrary, any reasonable norms for informed consent will embody a set of shared expectations about how such cases are to be interpreted, and not only so as to accommodate the semantic explosives of the philosophical logician (e.g. by factoring

in time-constraints, unforeseen events, or institutional protocol). Finally, I will mainly treat “absence of consent” as neutral between consent not being elicited and consent being elicited but refused. Thus, the students may decide to raid my books without asking me, or in spite of my insistence that they are not for them to borrow. This is not to say that there is no ethical difference between the two cases. Clearly there is (e.g. depending on whether I am readily available to be asked), but this is not a difference that will make the difference in the cases in which I am primarily interested here.

Second, what I say in what follows is premised on the understanding that the vast majority of interactions where we think that consent is ethically called for do not take place in the context of complex institutions such as universities or hospitals. On the contrary, most of these transactions take place in more informal settings, as when someone takes me along to the forest in order to teach me how to ski; then undertakes to bandage my subsequently sprained ankle; then eventually offers me a ride home. Much of the systematic thinking that has been produced in recent decades about the ethics of consent has focused on the need to articulate reasonable and enforceable guidelines for application in specific institutions (such as hospitals) in response to particular events (such as cases of mistreatment) located in particular contexts of intervention (such as the removal and use of human tissue, or a rising culture of litigation). All this systematic thinking serves a specific set of societal needs to establish mutually recognised and reasonably precise “criteria”, such as legal standards or codes of good practice. At the same time, when articulating a systematic rationale for these practices, it is reasonable to appeal to values and aspirations that have a history outside of these particular contexts. These facts together produce the following tension. On the one hand, the systematic theoretical rationale (e.g. respect for autonomy) provided for a given practice (e.g. informed consent requirements) may come to look excessively abstract or “idealistic” to the

people (e.g. mental health professionals) who are faced with the task of applying it in the context of day-to-day care for individual patients; each and every one of whom is always going to be truly “unique” in some particular way. On the other hand, the implementation of this rationale by those who are tasked with doing so can come across as narrow-minded, bureaucratic, “commodifying” or self-serving to other interested parties (such as patient representatives, or university academics who analyse the institutional administrative apparatus at work). In order to resolve this tension, both sides may have to give. The structural incentives and constraints that drive discussions of consent requirements in the context of specific institutions ultimately gain their legitimacy from values that are not only present in society more widely, but that will frequently pre-date and survive those institutions. Hence, we are entitled to think about the value of consent, for example, outside these institutional contexts in order to inform our views about how to respect it within them. At the same time, our commitment to the value in question derives its legitimacy at least in part from our ability to realise it within the institutional contexts where it actually needs to be applied “in real time”. Hence, we are entitled to complain when systematic theoretical justifications of a given practice (such as informed consent requirements) fail to speak to the decisions that institutions, clients, patients and their representatives actually face. We need to recognize that there is nothing intrinsically wrong about the fact that people in paid employment should want to cover themselves against complaints or litigation by asking service users to certify that they actually agree to central aspects of what is going to happen to them. In the absence of such safeguards, not only would much professional work involving substantial risk of harm to others be prudentially reckless. By effectively absorbing the decision-making role of all the patients in their care, the moral burdens imposed on carers would be potentially overwhelming.<sup>3</sup>

---

<sup>3</sup> I am grateful to Mohammed Abouelleil Rashed for discussion of this point.

Third, it is possible to value consent without thinking of it as either overriding, or some kind of master value. This is not only because there are other things (such as health outcomes or efficiency) that we could value more. It is also because prioritising consent can actually end up undermining the values we want it to serve (c.f. Groll 2012). There are at least four reasons for this. First, eliciting consent can actually be a manipulative form of behaviour. Thus, an unusually generous person who is highly anxious about saying “No” to requests can be dominated by others who pick up on this fact and use it opportunistically to gain advantage.<sup>4</sup> Second, refusing to give consent can equally be a manipulative form of behaviour. Thus, a person who seeks to dominate another can manipulate them by repeatedly refusing to consent to things they either have nothing against or, in the extreme case, actually want and are hoping to achieve.<sup>5</sup> Third, giving consent can itself be a manipulative form of behaviour. Thus, a person who is willing to do an unusually large number of favours in order to thereby gain status, power or protection may cultivate their disposition to give consent as part of a long-term strategy of making the people around them feel indebted, dependent on their good will, or subject to their control.<sup>6</sup> Fourth, some practices of eliciting and giving consent are superficial, meaningless, or even oppressive. As anyone knows who has even minimal experience of working in a complex modern institution (such as a university or a public health service), or who has ever done so much as to purchase an item online (such as a standard plane ticket), some legally enforceable practices of eliciting consent arguably have much less to do with respecting the autonomy of anyone involved than “going through the

---

<sup>4</sup> This point should be familiar to anyone who has ever been a victim of aggressive sales tactics.

<sup>5</sup> There are notorious issues I pass over here about the distinction between cases where “No” means No”, and cases where someone is “playing hard to get”, or simply playing.

<sup>6</sup> There are complex issues here I pass over about how “generosity” can function as a form of entrapment. One obvious, if imperfect, analogy is between being very cooperative on the one hand, and offering gifts or bribes on the other.

motions” in a highly asymmetric relationship that leaves very little room for the pursuit of acceptable alternatives.<sup>7</sup> Thus, it is a well-worn criticism of capitalist liberal democracies that while giving people the “illusion” of multiple freedoms, they are simultaneously in the business of perpetuating a battery of “legitimation myths” for practices that are said to remove from the vast majority of the individuals affected many forms of meaningful choice (see e.g. Zizek 2014). Be that as it may, the people who make such criticisms are not seriously suggesting that the case for consent requirements is, at all times and in all places (such as in your average transaction with a hospital nurse), conditional on this aim. Yet awareness of situations in which consent requirements actually undermine the values they are meant to promote is an indispensable constraint on any plausible account of its ethical significance, whether in theory or in practice.

### **3. Consent in the “non-ideal” case**

In some cases where consent is sought it is not only true that the person asked for consent fails to exhibit the standard marks of autonomous, or self-governing, rational agency. It is also true that the person in question, when faced with a decision, is likely to fail to choose what is in their best interest; will forget or regret that they have made the decision in question; will hold it against the person or persons eliciting consent that things have turned out as they have as a result of that decision; or will not fully understand the significance of their consent or what they have consented to. Let’s call a case that meets these conditions a

---

<sup>7</sup> Of course, we are usually “granted” a number of protections in return, of varying degrees of credibility.

“non-ideal case”.<sup>8</sup> Why be concerned to elicit the consent of such a far from fully autonomous agent, whose choices you know are likely to be an unreliable, or poor, guide to what is in their own best interest? What, if anything could be the legitimate source of complaint if you disregard the question of someone’s consent in such a “non-ideal” case?

There is a range of familiar answers to this question; none of which will cover all possible cases, but each of which will cover some (c.f. Scoccia 1990; Groll 2012; Lillehammer 2012). A first range of answers is unlikely to speak directly to the main problem addressed in this paper, but is nevertheless worth noting for the record. First, in failing to get my consent you may cause me displeasure or frustration, or even make me hard or impossible to deal with; and hence to this extent be acting against my (or even your own) best interest. By getting my consent you could simply be making sure that we are able to get along. Second, in failing to get my consent you may be acting against a general expectation in our group (or in society as a whole); thereby causing displeasure or frustration; hence potentially making both me and others even harder to deal with in the future; hence be acting against a general, or shared interest. Third, in failing to get my consent you may express a lack of respect for *people like me*; thereby showing *yourself* in a bad light; hence displaying a personal character defect or vice – regardless of whether you are acting in my best interest or not. Fourth, in failing to get my consent you may fail to show due respect *for the person I once was* (e.g. if I am suffering from some degenerative disorder) or *the person I could have been* (e.g. if I am suffering from a permanent disability). Fifth, in failing to get my consent you may fail to show respect for the fact that I am a member of some ethically salient category, where that category is not defined specifically in terms of a capacity for autonomous agency. Thus, I could respond to

---

<sup>8</sup> As previously noted, this will include a range of cases in which patients are deemed to have “capacity” or “competence”.

your failure to get my consent that you are refusing to treat me as a *family member, a fellow national, or a fellow human being* (c.f. Williams 2006).<sup>9</sup> All of this is true, but fails to speak fully to the idea that there is something missing in failing to elicit *consent in particular* that is explicable in terms of the specific individual I actually am, here and now, in at least some of the “non-ideal” cases at issue in this paper.

A second set of answers is able to capture at least some of these “non-ideal” cases head-on, but is unable to capture other cases of the kind. First, failure to elicit consent could be a missed opportunity to help me acquire or enhance a capacity to exercise fully autonomous, or self-governing, rational agency. Its wrongness could therefore be analogous to the wrongness of refusing someone therapy or treatment, or failing to provide education to a person with special needs (c.f. Pickard 2013). This could explain consent requirements in cases where either improvement or recovery is possible, and also in a wide range of cases beyond that, at least insofar as we factor in the uncertainty of outcomes. It does not, however, explain consent requirements in cases where the probability of improvement or recovery approaches zero (such as in cases of advancing dementia). Second, a failure to elicit consent could be a missed opportunity to stop, or slow down, a process of losing the capacity for fully autonomous, or self-governing, rational agency. Its wrongness could therefore be analogous to the wrongness of refusing someone preventive or controlling treatment or medication. This

---

<sup>9</sup> Another alternative is to explicate the respect I am owed in terms of the fact (if it is a fact) that I am a *person*. Unfortunately, the moral category of “personhood” is so contested with respect to the question of whether, and to what extent, moral personhood requires the capacity for autonomous agency that it is likely to be of limited use in the present context. Here is Harry Frankfurt: “To be a person, as distinct from simply a human organism, requires a complex volitional structure involving reflective self-evaluation. Human beings that lack this structure may be free of inherent volitional conflict, but they are not persons.” (Frankfurt 1999, 103) It might be asked if the domain of agents at issue in this paper is actually exhausted by the category of persons, however that notion is best understood. I do not propose to answer this question here. I am grateful to Neil Manson for raising the issue in discussion.

could explain consent requirements in cases where, even if improvement or recovery is not a possibility, there is some likelihood of a person being able to live and cope with their condition as long as possible, or longer than they otherwise would. It does not, however, explain consent requirements where the probability of slowing down the process of deterioration approaches zero (such as in cases of permanent incapacity). Third, you might ask for my consent on the fictional assumption, or pretence, that I actually have these capacities. Perhaps there are cases of therapy, treatment or education as just described that approximate to this case (As in “Let’s pretend to be adults”). Partly for that reason, and partly because I am looking for ways to capture consent requirements that don’t crucially depend on insincerity, delusion or make-believe, I will not discuss this possibility further. This is not to deny that it can sometimes be justified to approach a particular case in a way that abstracts to some degree from the established facts and probabilities in play. On the contrary, fictions, pretence and make-believe arguably play a more important role in the psychosocial dynamics of such cases than standard philosophical treatments have traditionally been prone to give them.<sup>10</sup>

So what could possibly be the point of giving someone a choice to consent in conditions where there is no expectation that they will successfully exercise fully autonomous, or self-governing, rational agency in the consideration, making, or execution of that choice? Consider some parallel cases. First, it is often appropriate to ask someone for their opinion even when the final decision will actually be made by someone else. Being invited to give an opinion (i.e. being consulted) is obviously not the same as being given the opportunity to refuse, yet some invitations to refuse actually function more like invitations to have an

---

<sup>10</sup> In much the same way, such treatments have normally tended to stay well clear of the interpretation of the symbols, myths and rituals that inform the practice of common forms of health care.

opinion even if they are couched in the language of consent (as in “Do you mind if I... pour you a glass of water; push ahead; make your bed; schedule the exam...”). Hence, just as there are offers that are meant to be refused (as in: “Would you like me to... take the class for you; make your clothes; never contact you again...”), there are offers to refuse that are meant to be refused (As in “If you really don’t want to...”). Indeed, the class of offers to refuse that are meant to be refused include both cases where the refusal to give consent is expected to lead to inaction (as in: “No, not now. I’m too busy”) and where it is expected to escalate the situation or take it to another level (as in “OK, I will be back later”; or “Well, I’m sorry; but we will need to... give you food; receive your essay by the deadline; make sure that all your debts are paid”). Furthermore, there are offers to refuse that are not meant to be refused as such but the failure to refuse of which can function to slow down or otherwise restructure the project of realising a specified outcome the ultimate shape of which is not assumed to be fully within the control of the party being asked for consent. (Some complex medical procedures may fall into this category.) Indeed, the fact that the outcome of the transaction is in some sense pre-determined, or implicitly assumed, does not remove the rationale of asking for consent. We can see the rational acceptance of this fact in a wide range of practices that, if considered purely in the abstract, may seem curious or even bizarre but, when reflecting on their social significance, are far from obviously crazy. To take just one topical example of the kind from the political arena: some institutional checks and balances function in an analogous way, even in the face of (or, perhaps precisely because of) significant asymmetries in power and understanding between the balancing parties. Thus, the parliament of a constitutional monarchy may be legally obliged to subject its major legislative decisions to royal, or some kind of other constitutional, approval (such as a referendum). As we all know, this does not always mean that the consenting “agent” in question is seriously expected to refuse, or that if he or she is reluctant to agree the proposals will not be carried out in some form anyway;

either immediately, after some delay, or after further discussion or amendment (e.g. where a referendum is treated as consultative). Nor is it always the case that the “agent” being asked to consent is either fully on top of what the decision involves; has the capacity to work it out for him or herself; has the power or ability to execute the decision; or will not come to regret the decision once it is finally made. (Imagine the – historically quite realistic - case where the “constitutional monarch” a minor.) All this notwithstanding, sincerely subjecting a proposal to said “agent” for their consent can embody a distinctive kind of respect for said “agent”, however peculiar that respect may seem from the point of view of standard discussions of contractual agreements between “equal” and “autonomous” parties. If things were always otherwise, the path from knowing “how the system works” to the conclusion that the system is absurd would be very short indeed. And still, at least for many of us, it isn’t. Different systems of checks and balances can have a complex and historically conditioned rationale, and can be implemented in different ways consistently with the insight that their function is to underwrite the legitimacy of mutually significant decisions in which it matters that the relevant participants in some sense *make them together*, and where giving the participants an opportunity to play a part in decisions that affect them is an important way of showing respect for the individual agents they actually are.

Of course, the political theory of “the consenting public” has often been taken to presuppose a conception of the individuals asked for consent as fully autonomous, or self-governing, rational agents (although quite a bit of dubious idealization is normally required even there. See e.g. Rawls 1971.) This is a condition that, by hypothesis, is not met by the consenting individuals in the kind of “non-ideal” case at issue in this paper. Nevertheless, my suggestion is that an analogous rationale can be provided for consent requirements involving people who fall short of exhibiting the standard marks of fully autonomous, or self-governing, rational

agency. On this view, the interventions the persons in question can reasonably be expected to be offered the opportunity for consent to are interventions (e.g. of a medical kind) that will primarily affect themselves, as well as those who care for, or otherwise interact with, them. They involve decisions that benefit from an analogous system of “checks and balances”, even if in certain cases the probative force of the refusal to consent only gives rise to *prima facie* obligations, or is otherwise defeasible to the point of being reasonably overridden by other concerns. To this extent, the process of eliciting consent can be said to have not only *instrumental* (as in avoiding patient frustration and resistance), or *epistemic* (as in getting all relevant information about the patient) value; it can also be said to have *intrinsic* value, in virtue of respecting the fact that the decisions in question are ones in which both the agents asking for consent and the agents being asked for it should be, in some non-trivial sense, *meaningfully involved*.<sup>11</sup>

#### **4. Consent, meaning and agency based interests**

What, if anything, can be said to explain or justify the claim that agents in a “non-ideal” case should be meaningfully involved in decisions that affect them? The present proposal is that having the opportunity to be meaningfully involved in decisions that principally affect us is something in which we have an *interest*, and hence that not to be given an opportunity to give or withhold consent to those decisions is to frustrate that interest by excluding us from the

---

<sup>11</sup> People doing things together in the sense relevant here is likely to depend on the establishment and maintenance of non-trivial levels of trust. The application of consent requirements can be one way of promoting or protecting such trust. I am grateful to an anonymous referee for pointing out the connection between consent and trust in this connection. A discussion of the nature of trust, and whether its role in consent is one that should be thought of as having either intrinsic or extrinsic value, would take me too far afield in the context of the present paper.

collective that normally should be involved in making them. What explains that we have this interest is the fact that we are agents capable of interactions of the requisite kind (e.g. interactions involving the eliciting of our consent with respect to what is to happen to us). This *agency-based* interest is an interest that someone can possess without themselves being able to effectively pursue it; protect it; defend it; or even fully articulate either it or its ethical rationale. Yet it is an interest that agents who fail to display the marks of fully autonomous, or self-governing, rational agency can share with agents who do display those marks, and therefore an interest that can rationalise consent requirements for agents in both categories. In other words, it is an interest that can ground a set of reasons to elicit consent that all the agents in question (autonomous or not) in some sense “share” (although not in the sense that some Kantians would say that there is a distinctive class of reasons that fully autonomous, or self-governing, rational agents all “share” by means of their capacity to autonomously, or otherwise rationally, grasp them). This agency-based interest therefore falls into the same broad category as other interests attributed to vulnerable others and that are sometimes appealed to in order to make sense of ethical constraints in our dealings with frail or disabled adults, children, or some non-human agents.

There are at least two different ways in which we can understand the idea of an agency-based interest. On the first (and narrower) interpretation, agency-based interests exist in virtue of some specific exercise of agency on the part of its holder, and are in that sense “self-generated”. Thus, it could be said that I have an interest in having have an office with enough space for books in virtue of once having decided to become a university academic. On the second (and wider) interpretation, agency-based interests exist in virtue of their holders having the property of being able to exercise agency (or being capable of exercising agency in certain ways, such as making decisions along with others). In that sense, agency-based

interest are not necessarily self-generated because something can be good for someone merely in virtue of their being able to exercise agency (or being capable of exercising agency in certain ways), and so independently of whether, and if so in what particular way, that capacity for agency is actually exercised on any occasion. Thus, it could be said that I have an interest in maintaining basic physical coordination even if I am currently too maimed or incapacitated to actually consider the issue.

It is the latter (and thus wider) interpretation of agency-based interests that underpins the argument of this paper. The significance of this is as follows. First, it is not being assumed that the interests providing a rationale for consent requirements are a function of the particular way that someone's agency *has actually been exercised* on any occasion. Thus, it is an interest I can in principle have even if it is one that I take no interest in, or have even renounced. (In the latter case, it may be natural to think that any *prima facie* obligations to which it gives rise would sometimes be overridden.) Second, it is not being assumed that the interests providing a rationale for consent requirements are a function of some particular way the individual possessing them is capable of exercising them. Thus, it is an interest I can in principle have even if it is one that I have no effective control over myself, or that I cannot even understand very well (if at all). It follows that agency-based interests (thus widely understood) could in principle be attributed even to agents who fail to meet standard criteria of capacity or competence. Third, insofar as it might - quite reasonably - be thought that agency-based interests are ethically *more significant* to the extent that their existence and content *does* depend on some particular actual exercise of agency, this could explain why basic capacity or competence is often thought of as a threshold above which consent considerations carry a distinctively important, or even decisive, ethical force. This does not, however, imply that such considerations carry no ethical force at all beyond the domain of

competence or capacity. Indeed, it is one of the main advantages of the argument of this paper that by appealing to the latter (and so wider) interpretation of agency-based interests we can make sense of the claim that they sometimes do, and why.

As previously noted the agency-based interest respected by eliciting consent in a “non-ideal” case is an interest that can be realized consistently with its exercise being such as to issue in choices, actions and outcomes that are otherwise *not* in the agent’s interest, much less in their *best* interest, or in their interest *overall* (as stipulated in the definition of a “non-ideal case”). It follows that the argument for respecting consent in the “non-ideal” case is not the standard kind of argument in the literature that we should respect the choices of agents because this is likely to be instrumental to, or a realization of, what is in their best interest (c.f. Dworkin 1994). On the contrary, it is an argument that preserves both the claim that in failing to elicit consent in a “non-ideal” case you can *fail* to respect someone’s interest, and the claim that in failing to elicit their consent can *promote* their best interest. The point is that in these cases you could be doing something *prima facie wrong* to that someone by acting in their best interest, even if they fail to exhibit the marks of fully autonomous, or self-governing, rational agency. If the interest you thereby fail to respect is such as to justify the imposition of *prima facie duties* on others, as might be argued on a so-called “Interest Theory” of rights, then the agency-based interest involved in decisions affecting oneself can in principle be thought of as a ground of duties on others to elicit consent, and thereby also of a *right* (however defeasible)

to be asked for and potentially refuse it.<sup>12</sup> Be that as it may, an Interest Theory of rights, interpreted consistently with the wider interpretation of agency-based interests introduced above, can explain the existence of any *prima facie* rights and duties we might wish to assert with respect to consent requirements in a “non-ideal” case in a way that is not correspondingly available on accounts of rights and duties grounded in the capacity for fully autonomous, or self-governing, rational agency (such as standard versions of a so-called “Will”, or “Choice”, theory), or some theory of rights grounding their existence in agency-based interests on the narrower interpretation introduced above. Whether this implication is an advantage or a disadvantage of an Interest Theory of Rights thus understood is a topic for another occasion.<sup>13</sup>

One advantage of the proposal just outlined is that it explains the ethical significance of consent in “non-ideal” cases *without* implausibly assuming that in order for failure to get my consent to count as problematic I must be thought of as either having exercised my agency in a given way, or as be in the possession of capacities I don’t actually have. According to this proposal, the justification of consent requirements in these “non-ideal” cases relies partly on the fact that although the person from whom consent is elicited does not possess all the features associated with fully autonomous, or self-governing, rational agency, they do possess

---

<sup>12</sup> The topic of the relation between rights and interests is complex and controversial. For a discussion broadly along the lines suggested in the main text, see e.g. Raz 1986. For an account of interests (e.g. in being able to give and receive promises; to make and break contracts, and potentially to give and receive consent) that I take to be consistent in principle with the account in the main text, see Owens 2012. Owens does not apply his account of what he calls “normative” interests to the range of non-ideal cases under consideration here, so it unclear whether he would extend their scope to cover cases where the putative holder of the relevant interest is effectively unable to fully understand or control its exercise.

<sup>13</sup> I have previously discussed this issue, but much too briefly, in Lillehammer 2012.

at least some of them; in particular certain features that make it possible to engage in meaningful activity, including some form of meaningful communication with others (the notion of “meaningfulness” in play does not imply that the activities in question are necessarily regarded *as* meaningful to all the persons involved at the time, nor – given their condition – that they could in principle be so regarded by them) The category of agents in question includes a wide range of persons suffering from chronic mental disabilities; developmental conditions; advancing dementia; anxiety disorders; addiction, and the like. For lack of a better label, I refer to individuals in possession of these features as having the capacity for *meaningful agency*. The present proposal is that this is a capacity that individual agents in a “non-ideal” case could actually have, and in terms of which any requirement to elicit their consent could potentially be explained or justified.<sup>14</sup>

The category of “meaningful agency” at issue here is intended to include agency that satisfies standard definitions of “capacity” or “competence”, without such definitions. Instead, the point of invoking the language of “meaningful agency” is to provide an informal conceptual currency which the ethical significance of legal definitions of “capacity” or “competence” and the like can be ethically discussed. One of the things the appeal to meaningful agency may help to provide is an answer to the question why we should care about a patient’s mere capacity or competence (however defined) when we have every reason to believe that its exercise is unlikely to be in the patient’s best interest. While I do not have a comprehensive theory to offer about what meaningful agency would consist in across all possible contexts,

---

<sup>14</sup> While the assumption that any genuine, or ultimate, reasons to elicit consent must be explained in terms of features *intrinsic* to the agents in question is generally accepted in the literature on this topic, it is obviously an assumption that might be questioned (for example in some of the cases provided in Section 3 above). Even if we relax that assumption, however, it remains a virtue of the present proposal that it is in principle able to explain the reasons in question by appeal to intrinsic features of agents. I do not, therefore, propose to discuss this issue further here.

I'm not really sure that one is needed for present purposes. I have three reasons for making this claim. First, there are indisputable cases of agents falling outside the domain of fully autonomous, or self-governing, rational agency for whom the question of consent at issue in this paper clearly does arise, and I have identified a number of such cases in the text. Second, there are indisputable cases of entities falling outside this domain, such as my pot plants or stereo, with respect to which this question clearly does not arise; or at least not in the same way. (There is obviously going to be significant scope for vagueness here, and there is no *a priori* reason to think that any reasonable "sharpening" would be either "speciesist", or restricted to organic life forms alone). Third, and with respect to the theoretical purposes at issue in this paper, there is no reason why a broadly circular and heuristic definition would not do (such as – and this is only a schematic example – "X is capable of meaningful agency *iff* some agent (be that X or someone else) can take up a "second-personal stance" with respect to X without irrationality, false belief, or self-deception").<sup>15</sup> The operational constraints imposed by a context where institutional norms and legal frameworks are to be formulated are obviously a different matter, but even here there is likely to be "traces" of broad circularity at work (e.g. in the interpretation of "capacity" or "competence" cited in Footnote 2 above).

In previous work, I have distinguished between respecting someone's *agent autonomy* on the one hand,<sup>16</sup> and respecting their *choice autonomy* on the other.<sup>17</sup> The notion of *meaningful*

---

<sup>15</sup> For a comprehensive discussion of the second-personal standpoint and its ethical significance, see e.g. Darwall 2006. Nothing in the main text should be read so as to imply the endorsement of Darwall's account of that notion in his book, which, for reasons that should be clear from the main text, would be excessively demanding for the purposes of the present paper.

<sup>16</sup> As defined in Lillehammer (2012), agent autonomy requires possession of the following marks of rational self-governance: i) manifestation of a capacity for higher order reflection and endorsement of practical options; ii) the actual manifestation of a capacity for planning and executing actions that accord with the practical options endorsed; iii) responsiveness of

*agency* identified in the previous paragraph is meant to give further content to the idea, proposed in that work, that there is a range of agents (not necessarily a single, or unique kind) for whom to respect *choice autonomy* is an ethical default option even in the absence of *agent autonomy*. The idea, to put it in summary form, is this. There is a form of respect we can reasonably be said to owe to some agents who are not capable of fully autonomous, or self-governing, rational agency in virtue of the fact that they are able to engage in communicative interactions that either they themselves, or those who interact with them (e.g. friends, family, carers, medical personnel, etc.), can reasonably think of as meaningful. This is a form of respect that can be owed to such agents irrespective of how their ability to engage in meaningful activity is related to their (non-existent, failing, or yet to be realised) capacity to exercise fully autonomous, or self-governing, rational agency; although the case for respecting their meaningful activity could obviously be enhanced (and radically so) by its relationship to some (possibly non-actualized) capacity to exercise fully autonomous, or self-governing, rational agency as well.

## References

Archard, D. 2008. Informed Consent: Autonomy and Self-Ownership. *Journal of Applied Philosophy* 25: 19-34.

---

reflection, endorsement and execution of options to minimally intelligible standards of rational argument; and iv) reflection, endorsement and execution of practical options that involve a conception of oneself as a single person living a certain kind of life.

<sup>17</sup> As defined in Lillehammer (2012), choice autonomy requires *at least* being presented with a non-empty set of (in the relevant circumstances) reasonable options between which one is able to choose, or otherwise express a preference. (C.f. the UK definition of “capacity” in Footnote 2 above.)

Alzheimer's Society (GB). 2018. Decision-making and Mental Capacity. Online. <https://www.alzheimers.org.uk/get-support/legal-financial/decision-making-and-mental-capacity#content-start> (accessed July 25 2018).

Arpaly, N. 2003. *Unprincipled Virtue*. Oxford: Oxford University Press.

Beauchamp, T. L. 2005. Who Deserves Autonomy, and Whose Autonomy Deserves Respect? In: *Personal Autonomy: New Essays on Personal Autonomy and its Role in Contemporary Moral Philosophy* (pp. 310-325), J. Stacey Taylor (ed.) Cambridge: Cambridge University Press.

Beauchamp, T. L. and Childress, J. F. 2008. *Principles of Biomedical Ethics*. 6<sup>th</sup> Edition. Oxford: Oxford University Press.

Darwall, S. 2006. *The Second Person Standpoint: Morality, Respect and Accountability*. Cambridge: Harvard University Press.

Dworkin, G. 1988. Autonomy and Informed Consent. In: *The Theory and Practice of Autonomy* (pp. 100-120), G. Dworkin. Cambridge: Cambridge University Press.

Dworkin, R. 1994. *Life's Dominion*. New York: Vintage.

Elster, J. 1984. *Ulysses and the Sirens*. Cambridge: Cambridge University Press.

Estlund, D. 2007. *Democratic Authority: A Philosophical Framework*. Princeton: Princeton University Press.

Eyal, N. 2011. Informed Consent, *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/informed-consent/> (accessed February 6, 2017).

Feinberg, J. 1986. *Harm to Self*. Oxford: Oxford University Press.

Frankfurt, H. 1999. The Faintest Passion. In: *Necessity, Volition and Love* (pp. 95-107), H Frankfurt. Cambridge: Cambridge University Press.

Gunderson, M. 1990. Justifying a Principle of Informed Consent: A Case Study in Autonomy-Based Ethics. *Public Affairs Quarterly* 4: 249-265.

Groll, D. 2012. Paternalism, Respect, and the Will. *Ethics* 122: 692-720.

Levy, N. 2006. Autonomy and Addiction. *Canadian Journal of Philosophy* 36: 427-447.

Lillehammer, H. 2012. Autonomy, Value and the First Person. In: *Autonomy and Mental Disorder* (pp. 192-213), L. Radoilska (ed.), Oxford: Oxford University Press.

Kittay, E. 2005. At the Margins of Moral Personhood. *Ethics* 116: 100-131.

Korsgaard, C. 2009. *Self-Constitution*. Oxford: Oxford University Press.

Manson, N. C. and O'Neill, O. 2007. *Rethinking Informed Consent*. Cambridge: Cambridge University Press.

McMahan, J. 1996. Cognitive Disability, Misfortune and Justice. *Philosophy and Public Affairs* 25: 3-35.

Miller, F. G. and Wertheimer, A. (eds.), 2010. *The Ethics of Consent*. Oxford: Oxford University Press.

Nussbaum, M. 2009. The Capabilities of People with Cognitive Disabilities. *Metaphilosophy* 40: 331-351.

O'Neill, O. 2002. *Autonomy and Trust in Bioethics*. Cambridge: Cambridge University Press.

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

Pickard, H. 2013. Responsibility without Blame: Philosophical Reflections on Clinical Practice. In: *The Oxford Handbook of Philosophy and Psychiatry* (pp. 1134-1154),

K. W. M. Fulford et. al. (eds.). Oxford: Oxford University Press.

Radoilska, L. 2015. Autonomy in Psychiatric Ethics. In: *The Oxford Handbook of Psychiatric Ethics* (pp. 354-371), J. Sadler et. al. (eds.). Oxford: Oxford University Press.

Rawls, J. 1971. *A Theory of Justice*. Oxford: Oxford University Press.

Raz, J. 1986. *The Morality of Freedom*. Oxford: Oxford University Press.

Savulescu, J. 1994. Rational Desires and the Limitation of Life-sustaining treatment. *Bioethics* 8: 191-222.

Scoccia, D. 1990. Paternalism and Respect for Autonomy. *Ethics* 100: 318-334.

Sunstein, C. and Thaler, R. H. 2008. *Nudge: Improving Decisions about Health, Wealth and Happiness*. New Haven: Yale University Press.

Velleman, J. D. 1992. Against the Right to Die. *Journal of Medicine and Philosophy* 17: 665-681.

Williams, B. 2006. The Human Prejudice. In *Philosophy as a Human Discipline* (pp. 135-152), A. Moore (ed.), Princeton: Princeton University Press.

Zizek, S. 2014. *Trouble in Paradise: From the End of History to the End of Capitalism*. London: Allen Lane.