

## **Toward an Agent-Centered Theory of Voluntariness\***

**Abstract:** The concept of voluntariness is central to informed consent and personal autonomy, yet it has been underexplored by bioethicists. There are various theories intended to explain voluntary choice and action. None is fully agent-centered, in the sense that the conceptualization of voluntariness takes into account the agent's views of her decisions and actions. An agent-centered theory of voluntariness would promote analytical precision, and foster autonomy in healthcare and research practices. According to the most influential bioethical theory of voluntariness, here called the Voluntariness as Control theory, an action is non-voluntary if the agent is controlled by external influences. The theory is critically discussed from an agent-centered perspective, and a new conceptualization of voluntariness is proposed.

**Keywords:** voluntariness, autonomy, informed consent, agent-centered, bioethics

### **Introduction**

Voluntariness is a central concept in bioethics. It is most commonly associated with informed consent, i.e., an agent's valid authorization of a medical intervention or research participation. Consent to such interventions or practices is invalid if it is not voluntary. However, the concept of voluntariness has been underexplored by bioethicists. Multiple and mutually incompatible theoretical accounts are present in the literature on informed consent. Although one type of conceptualization has been more influential than others, no single tradition or school of thought overshadows the others completely. This article is intended to stimulate bioethicist debate on voluntariness; its conditions, and role.

No theory of voluntariness that is present in the contemporary bioethical literature is agent-centered, i.e., takes the agent's perspective into account. Here, it is argued

---

\* Submitted for review.

that, to foster personal autonomy, a conceptualization of voluntariness should be aligned with how the agent actually leads her life and acknowledge whether she is engaged with her decisions and actions. It is proposed that an action should be understood as non-voluntary if and only if it is coerced, and if the agent thinks of the action as non-voluntary when being coerced to perform it.

The article is structured as follows. First, I introduce three hypothetical cases that demonstrate the ambiguities of voluntariness as a pre-theoretical construct. Then, I outline five theories of voluntariness that are present in contemporary bioethical literature. The most influential, here called the Voluntariness as Control theory, is the main target of the subsequent arguments. In the sections that follow, I use a set of examples to show how the Voluntariness as Control theory is flawed, and why it is important that a theory of voluntariness is agent-centered. Lastly, I defend my proposed conceptualization from some expected objections, and put the theory at work by applying it to hypothetical cases.

## **Contemporary theories of voluntariness**

### **When voluntariness is relevant**

Consider these three hypothetical cases in which the concept of voluntariness could be invoked. First, suppose that a nurse serves a patient a cup of coffee. The patient, who does not wish to take any medicine, asks whether there is anything but coffee in the cup. It is only coffee, the nurse assures the patient, knowing that she does not wish to take any medicine. This continues routinely for a month or two. The nurse serves the patient coffee, persistently maintaining that there is nothing but coffee in the cup, and the patient consumes it. One day, the patient learns that there has been medicine dissolved in every cup of coffee that the nurse has served her. Did the patient take the medicine voluntarily?

Now, suppose that a young pregnant woman tells her doctor that her parents, her partner, her partner's parents, and all her friends think that she should have an abortion. She quotes them saying: "You are too young. You should finish school, get a job, marry, and then perhaps think about having children." When her doctor asks the patient what *she* wants, she repeats in first person: "I am too young. I should finish school, get a job, marry, and then perhaps think about having children." It is

clear to the doctor that the patient does not follow her own wishes, but others.<sup>1</sup> If she would go through the procedure, would it be voluntarily?

Lastly, suppose that a patient has lived with a severe depression for a long time. She also suffers from some physical disease, the cure of which requires that the patient undergoes treatment. The patient refuses, and expresses a wish to die. Does she refuse treatment voluntarily? If her doctors enrolled the patient in a study to learn about the process of dying from that disease, would the patient's consent to participate be valid?

In neither case is the answer obvious. They are—supposedly—examples of manipulation, social pressure, and the potential lack of decision-making capacity, which all require careful moral analysis. However, it is not unquestionably clear whether the concept of voluntariness should be invoked to express whatever it is that is morally problematic with the cases. Conceptual ambiguities render it vague how the concept of voluntariness should be applied in healthcare and research. In what follows, I will briefly account for some conceptualizations of voluntariness that have been proposed in bioethical literature. Some demarcations are necessary, however.

First, some theorists wish to distinguish between voluntary choice and voluntary action. However, in the bioethical literature on voluntariness, “choice” and “action” are more often than not used indiscriminately. And, I fail to see why a distinction would be necessary. In the present context, the concept of voluntariness—disregarding of how it is understood—applies to informed *consent*. Consent could also be conceptualized both as choice and as action, so that consenting to something is “to decide in favor of that something,” or similar, or “to perform a speech act that marks an approval of that something,” or similar. Voluntariness would apply alike in either way; the agent consents voluntarily or non-voluntarily, disregarding of whether consent is understood as choice or as action. Accordingly, “choosing” and “acting” voluntarily are here merely proxies for the overarching “consenting” voluntarily, and will be used interchangeably.

Furthermore, in the present context, voluntariness is an interpersonal concept (cf. Nozick 1974, p. 262). Since ancient times, voluntariness has been discussed in wider contexts. Aristotle, for instance, questioned whether a sailor who throws his cargo overboard to save his ship from capsizing in a storm does so voluntarily (Aristotle. 2014, 1110a5–10). Facts of nature that forces us to make choices like the

<sup>1</sup>Whether observers really can determine the authenticity of others' desires is a whole other issue (see Ahlin forthcoming).

sailor's can be analyzed in terms of voluntariness, although it is not relevant for the present interpersonal purposes.

Moreover, doing something *non-voluntarily* is different from doing something *involuntarily*. While it is yet to be established what the former means, doing something involuntarily means that the action is not causally linked to the agent's decisional mechanisms. In short, reflexes such as blinking and sneezing are involuntary actions. It could also be argued that they are not actions at all, but I will not explore that possibility further here. It suffices to say that in the following, involuntariness is not of interest, why I will only treat non-voluntariness (and voluntariness).

Finally, it has been argued that the individualist framework of personal autonomy and informed consent is inappropriate and should give way to a structuralist narrative that better captures agents' lived experiences by acknowledging patterns of "class, race, ethnicity, sex, and other demographic characteristics" (Fisher 2013, pp. 256–60).<sup>2</sup> Others have suggested that individualism "abstracts from an interconnected social reality" and "overlooks the fact the people often base the social relationships on relational care" (Osuji forthcoming). Such critique marks a distinction between individualism and non-individualism that, in turn, entails a complete separation between different schools of bioethical thought. Therefore, I note that a defense of individualism *per se* would be desirable, although that is beyond the scope of the present purposes. Non-individualist theories of voluntariness will not be taken into further consideration here.

### Theories of voluntariness

Most accounts of voluntariness share the notion that coercion always renders actions non-voluntary. Hyman discusses (but dismisses) an alternative view: "... it is plausible that threats, however terrible, change the incentives associated with a course of action, just as offers and bribes do, so that they *influence* but do not *abolish* choice, and therefore do not negate voluntariness" (Hyman 2013, p. 691). Olsaretti has developed a theory of voluntariness that does not build on coercion. In it, a choice is non-voluntary if it is made because the alternatives that the chooser believes she faces are unacceptable according to an objective standard of well-being (1998, 2008; see also Colburn 2008). That is, if *A* is coerced to do  $\phi$ , but still has the acceptable

<sup>2</sup>See also chapter 4 in Bowman et al. (2012) for an analysis of consent in medical settings resting on power structures and authority.

alternative of doing  $\delta$ , then she does  $\varphi$  voluntarily. Furthermore, Olsaretti holds that choice is voluntary if the only acceptable alternative is one that the agent likes so much that she chooses it because of that. So, if  $A$  is coerced to do  $\varphi$ , but likes  $\varphi$  so much that she chooses to do it because of that, then  $A$  does  $\varphi$  voluntarily.<sup>3</sup>

According to Hyman, the concept of voluntariness must be understood in light of its function, which is to “inform the appraisal of individual conduct,” in particular “the assessment of innocence and guilt” (2013, p. 685). Hyman suggests that voluntariness is formed by negation, and provides the rough definition that “a certain thing is done voluntarily if, and only if, it is *not* done out of ignorance or compulsion” (2013, p. 685). He tracks the historical roots of his theory back to the ancients, noting that ignorance is identified as an exculpation already in the *Iliad* (2013, p. 686).

Nozick proposes that non-rights-violating actions are voluntarily performed. The theory builds on the so-called “self-ownership thesis,” i.e., the normative proposition that every individual person has an exclusive right to her own body and mind that amounts to a property in her self (Nozick 1974, p. ix; see also Ch. 1–2). As full property rights apply in Nozick’s political libertarianism, the self-ownership thesis entails that no person is morally entitled to intervene with anyone else’s person or project. Intervening (without permission) is a rights violation. In Nozick’s theory, all rights-violating actions, and only rights-violating actions, negate voluntariness.

I call the theory which has been most influential in bioethics the Voluntariness as Control (VaC) theory. It is supported by, among others, Beauchamp and Childress (2013) and Nelson et al. (2011). According to the VaC theory, an action is voluntary if it is free from controlling influences. In the coffee example, the patient was manipulated into believing that there was no medicine in her coffee. She was therefore under someone else’s control, and acted non-voluntarily. Likewise, had the nurse pointed a gun at the patient and said “drink your coffee or I will shoot you,” the patient had been controlled by coercion and thus acted non-voluntarily. As the VaC theory is the most influential in bioethics, it will be the main target for the arguments in the following sections. Therefore, it deserves some elaboration.

### The VaC theory

In their prominent book *Principles of Biomedical Ethics*, Beauchamp and Childress adhere to the VaC theory. They outline their version of the theory by writing that “a

<sup>3</sup>See Wertheimer (2012) for a critical discussion of Olsaretti’s theory.

person acts voluntarily if he or she wills the action without being under the control of another person or condition” (2013, pp. 137–40). Similarly, Appelbaum et al. argue that voluntariness depends on “the extent to which subjects are actually exposed to external, intentional, and illegitimate influences that causally impact their decisions” (2009, p. 37). Nelson et al. (2011) provide the most elaborate account of the theory that voluntariness is closely linked to being in control over one’s actions.<sup>4</sup> Voluntary action, they argue, should be understood in terms of the two necessary and jointly sufficient conditions of intentional action and the absence of controlling influences. The notion of intention is binary, in the sense that an act either is or is not intentional, while the notion of controlling influences is a matter of degree, so that an act can be more or less free from controlling influences on a continuum from total control to total absence of control.<sup>5</sup>

Examples of controlling influences in the broad sense include offers of payment, threats, education, deceit, manipulative advertising, emotional appeals, and the like (ibid., p. 7). Such influences can deprive agents of at least some degree of voluntariness. Manipulation involves “the use of nonpersuasive means to alter a person’s understanding of a situation and motivate the person to do what the agent of influence intends” (ibid., p. 8). A person can be manipulated in several ways. One can manipulate the information a person receives through different communication techniques or the format and method of risk disclosure. Financial incentives such as offers or rewards or access to drugs or medical care can distort a person’s view of her options of choice. Furthermore, one can be manipulated through, for example, withheld information, misleading exaggeration, and explicit lies, which are all examples of cases in which the manipulated agent has no credible possibility of recognizing that she is receiving skewed information.<sup>6</sup>

Furthermore, a person can be controlled through coercion. Coercion is the total control over an agent’s actions that occurs “if and only if one person intentionally either forces another person or uses a credible and severe threat of harm to control

<sup>4</sup>Their account forms the basis of at least one empirical study of voluntariness; Miller et al. (2011).

<sup>5</sup>According to Nelson et al., the theory of voluntariness as degree of control was first introduced as a detailed construct by Wall (2001). However, Wall did not conceive the notion of voluntariness as control. Beauchamp and Childress had already written that the “primary meaning of ‘voluntariness’ is exercising choice free of coercion or other forms of controlling influence by other persons” in the second edition of their *Principles* (1983, p. 87).

<sup>6</sup>Similarly, a person may be persuaded into doing or believing something. However, Nelson et al. argue that persuasion is consistent with voluntariness. When persuaded, “a person believes something through the merit of reasons proposed” and is therefore not controlled (2011, p. 7).

another person” (Nelson et al. 2011, p. 7). True coercion by threat “requires that a credible and intended threat disrupts and reorders a person’s self-directed course of action” (ibid., p. 8).<sup>7</sup>

According to the VaC theory, the patient in the coffee example took her medicine non-voluntarily, at least given that someone—presumably the nurse—intentionally arranged the whole situation. If there had been a mix-up in the hospital kitchen so that the patient was served coffee intended to be served to some other patient who, for some reason, wished to have his or her medicine dissolved in coffee, the patient would not have been intentionally controlled by someone else. VaC theorists would then not deem the patient’s actions non-voluntary. However, the opposite conclusion could also be drawn, if a distinction is introduced between control *negatively* and control *positively*, in a sense corresponding to negative and positive liberty.

The above account of control as non-manipulation and non-coercion is an account of control negatively construed, that is, acting and choosing without being controlled by some other person or condition. Being in control positively is something else. Nelson et al. write that the condition of noncontrol “is not sufficient for voluntary action, because an agent *in control* is likewise essential,” and analyze “this agency” as “*capacity* for intentional action” (2011, p. 10). Similarly, Wall writes that control requires that one is “able to follow through and perform in ways that facilitate the realization of what is sought,” which appears to be the idea of being able to direct one’s behavior in accordance with one’s intentions (2001, pp. 130–1). This is also how Beauchamp indicates that he understands control (Beauchamp 2010, p. 69).

The patient who, due to a mix-up in the hospital kitchen, is mistakenly served coffee with medicine dissolved in it is not controlled by some other person or condition. Yet, she lacks control positively construed; she is not in control over her medication. Therefore, a VaC theorist could argue that she does not take the medicine voluntarily.

Control positively construed, however, shows too much resemblance to the concept of personal autonomy. It obscures the important distinction between doing something *voluntarily* and doing something *autonomously*.

Beauchamp and Childress describe autonomy “in terms of normal choosers who act (1) intentionally, (2) with understanding, and (3) without controlling influences that determine their action” (2013, pp. 102–4). Taylor writes that “a person is autonomous if he guides and directs his decisions and actions in the light of his own

---

<sup>7</sup>This conceptual framework was first introduced by Nozick (1969).

desires and values, free from the interference of others” (Taylor 2009, p. xiii). And, Christman understands it similarly:

In general, autonomy is meant to manifest self-government, the ability of the person to guide her life from her own perspective rather than be manipulated by others or be forced into a particular path by surreptitious or irresistible forces. (Christman 2009, p. 134)

Control positively construed is almost or completely identical to personal autonomy, as the concept is understood both generally in the field and specifically by some of the most prominent autonomy theorists. It is unclear what VaC theorists think of this resemblance. However, the distinction between voluntariness and autonomy should be upheld. It is important not least because voluntariness is partly constitutive of autonomy; an act can be voluntary and non-autonomous but not non-voluntary and autonomous. Therefore, I will not take the possibility of control positively construed into further consideration in this article.

## **An agent-centered theory of voluntariness**

### **Manipulation and voluntariness**

The VaC theorists’ claim that manipulation negates voluntariness is too strong. There might be a connection between the factual basis for a choice or an act and the voluntariness with which it is performed, but the VaC theory overstates that connection. The theory renders inaccurate act descriptions, and thereby fails to be aligned with how agents lead their lives. Consider the following hypothetical examples of manipulated agents whose choices it would be misguided to describe as non-voluntary.

Suppose that a patient correctly believes that she has treatment options *A*, *B*, and *C* to choose from. She is completely indifferent with regards to how preferable they are; it does not matter at all to the agent which treatment she receives. The hospital has financial interests in patients choosing *non-A*, and the patient’s doctor has been instructed to falsely tell patients that *A* is unavailable. The doctor complies, and based on the doctor’s lie the patient chooses *B*.

It is an overstatement to claim that the patient chose *B* non-voluntarily. The patient could live her whole life not knowing that she was once deprived of option *A*. In her mind, the choice stood between the equally good *B* and *C*. Describing her

treatment decision as non-voluntary is alien to how she leads her life. If the patient were to find out about the truth, she might begin to think of option *A* as somewhat better than *B*; an irrational but understandable psychological reaction, similar to what behavioral economists call “loss aversion” (cf. *Loss aversion*). Nonetheless, it would be inaccurate of her to suddenly begin to describe a treatment choice which she has always been satisfied with as non-voluntary.

Yet, VaC theorists would argue just that. The patient did not “will the action without being under the control of another person or condition,” as Beauchamp and Childress might wish to put it (see above). Likewise, Nelson et al. could point out that the patient did not believe that *B* is the best choice merely due to “the merit of reasons proposed,” but also due to the doctor’s lie, and is therefore choosing under controlling influences, i.e., non-voluntarily.

Their act description is uninformative—perhaps even deceptive, as it triggers certain intuitions that should not be triggered here. Certainly, even more so as the bioethicist moral vocabulary includes concepts such as “deceit,” “obligations,” and “rights,” which are more suitable than “voluntariness” to apply here, while being sufficiently accurate to explain the ethics of the case. Given that the patient was wronged in some way, the concept of voluntariness is not indispensable for a description of precisely how she was wronged. In this case, it suffices to say that the doctor lied to her, which is a breach of confidence and a restriction of the autonomy with which the patient made her treatment decision.

Consider now an example in which a patient is not robbed of an option. Instead, an option is added to her range of available choices, and the patient is lied to concerning the new option’s features.

A patient at a medical university correctly believes that she has treatment options *A*, *B*, and *C* to choose from. She ranks the alternatives in that order. Now, her doctor, who is also a researcher, informs the patient of an experimental treatment alternative *A+*. The doctor suggests that *A+* is superior to the other options, but the truth is that there is no evidence supporting the doctor’s claims. Based on the doctor’s lie, the patient chooses option *A+* and suffers painful consequences from medical maltreatment.

Again, VaC theorists would argue that the patient stood under the doctor’s control, and thereby chose non-voluntarily. However, at least while there are other relevant moral concepts available, it is an imprecise conviction.

Describing the patient’s choice as non-voluntary is alien to how she leads her life.

If she were to find about the truth, she would likely rank  $A^+$  far below  $A$ ,  $B$ , and  $C$ . Nonetheless, it would be inaccurate of her to suddenly begin to describe a treatment that she was misled to value incorrectly as non-voluntary first when, and because, she learns its true value.

Furthermore, the patient made what she thought of at the time as a conscious and self-determining choice. She was presented with a list of options, ranked them per her own preferences, decided in favor of one of them, and had her wishes adhered to. Describing her choice as non-voluntary is to belittle her role as a rational decision-maker, as the choice of terms indicates passivity, i.e., non-agency. It is to degrade the patient completely from person to prop on the scene of events.

Lastly, holding that the patient chose  $A^+$  non-voluntarily is inflationary. It deprives the words “non-voluntary” of an intensity that we should strive to retain to be able to vividly and accurately depict gun-to-the-head scenarios that are sometimes a reality.

It is out of place to say that the patients in these examples chose non-voluntarily. Doing so adds little of value to a description of their actions, particularly in relation to the precision that could be attained by applying the analytically superior concept of autonomy to the cases: The patients chose voluntarily but—for reasons unrelated to voluntariness—non-autonomously. A theory that renders the patients’ actions non-voluntary is suboptimal with regards to fostering personal autonomy and to analytical precision; it should be concluded that the notion that manipulation negates voluntariness is too strong.

### **Agent-centeredness**

I believe that grounding the concept of voluntariness *outside* of agents, as the VaC theory does, is inherently mistaken. The VaC theory essentially disregards the agent, as it is oriented around others’ control of her rather than being oriented around her directly; it is not agent-centered. But, voluntariness is deeply connected to individual agents’ engagement with their choices. Therefore, it is crucial that the agent’s subjective point of view is not neglected in a conceptualization of voluntariness.

My proposal is that an action should be understood as non-voluntary if and only if it is coerced, and if the agent thinks of the action as non-voluntary when being coerced to perform it. In the following sections, I will attempt to establish the view that the agent’s subjective point of view is important to a conceptualization of

voluntariness. Then, I will defend my proposal from three objections that I expect from critics.

Consider these examples, intended to demonstrate that the VaC theory disconnects the concept of voluntariness from agents' choosing and acting, and that it is problematic to do so.

Suppose (1) that a doctor informs her patient that she may choose between treatments *A*, *B*, and *C*. The hospital has financial interests in patients choosing between *A* and *B*, so the doctor has been instructed to manipulate the patient by purposefully depicting *C* as less favorable than it actually is and *A* and *B* as more favorable than they actually are; the doctor complies. Now, suppose (2) that the doctor does not purposefully depict *A*, *B*, and *C* as such, but that she mistakenly believes them to have the characteristics that she attributes to them in (1).

VaC theorists would argue that the patient is being controlled in (1) and therefore chooses non-voluntarily. However, in (2), the patient's decision is not being controlled and she is therefore choosing voluntarily. But, the patient made what she conceived of as two identical choices: She does not think of them as non-voluntary. Theories that give different answers to (1) and (2) neglect her point of view. They detach the concept of voluntariness from agents' choosing and acting.

Now, suppose (3) that the doctor tries to follow the hospital's instructions but becomes confused and, while believing that she is successfully deceiving the patient, she informs her about *A*, *B*, and *C* as they actually are. In (3), VaC theorists would again claim that the patient's choice is being controlled by the doctor, if poorly so. They would argue that the patient chose non-voluntarily, despite the fact that the doctor's intended manipulation had no effect on her choosing. The patient made the same choice as she would have made had she been treated by a non-manipulative doctor. VaC theorists, and others, must explain why a person's action *Y*, that has no effect on another person's action *Z*, renders *Z* non-voluntary disregarding of the *Z*-person's own beliefs. To be clear, the VaC theory has the absurd consequence that there could be situations when VaC theorists must try to convince people whose stable and considerate opinions are that they have acted voluntarily that they are mistaken: "Your views do not matter, you acted non-voluntarily."

An agent-centered theory could deal with these examples, as it would incorporate the agent's experience of choosing and acting. Therefore, the VaC theory is suboptimal in comparison to agent-centered theories in at least one respect.

**My proposal**

To repeat, I propose that an action should be understood as non-voluntary if and only if it is coerced, and if the agent thinks of the action as non-voluntary when being coerced to perform it. There are three main complexities with this conceptualization: It is not immediately clear what it is to think of an act as non-voluntary, why it should be thought when being coerced, or what function the notion of coercion has. They will be elaborated on in that order.

Not everyone has a stable and considered idea of what voluntariness is. Yet, my proposal seems to require just that, since the proposal is that agents must think of their actions as non-voluntary for them to be non-voluntary. What I have in mind though, is an agent's inner sense of resistance to the behavior that she is exposed to; an intolerance, anti-attitude, or unwillingness.

The inner sense must not be displayed by the agent, it suffices that she experiences it. An inner sense of resistance is displayed either verbally by articulating it, or physically by resisting. Neither should be required of agents for their actions to be labeled non-voluntary. Suppose, for instance, that a doctor points a gun to the head of a patient who refuses treatment with the intention of forcing her to undergo it. It is unreasonable to say that the patient then undergoes treatment voluntarily only because she did not protest or resist when the doctor pointed a gun at her. It is enough that the patient experiences a sense of inner resistance for the action to be non-voluntary, she must not also display it.<sup>8</sup>

The agent should think of the action as non-voluntary when being coerced to perform it. More precisely, it should be thought at least at some point while the coercion lasts. It could be the case that a person is coerced to undergo a long treatment during which her inner resistance slowly breaks down until she becomes emotionless or apathetic towards it. That does not mean that she continues the treatment voluntarily. But, if the person changes her attitude not due to the coercion, or to the influence of some other person or condition, but due to her own independent and rational reevaluation, it is not unreasonable to say that she continues her treatment voluntarily (although she begun it non-voluntarily). Furthermore, if an agent thinks of an action as non-voluntary first after the coercion has ended, it was not non-voluntarily performed. A conceptualization of voluntariness should not enable agents to dress up

---

<sup>8</sup>It is mainly through the notion of thinking of one's own actions as non-voluntary that the agent's perspective is taken into account by my proposal, i.e., what makes my proposal agent-centered.

their regret in an exculpatory language that downplays their own role in what has unfolded.

Finally, an action is non-voluntary only if it is coerced. Coercion is a form of influence not like any other. It alone bears the capacity to neutralize agency completely. Therefore, while other forms of influences can be gravely unjustified, they do not reach the same threshold as coercion does. Yet, I withhold that coercion does not negate voluntariness by default. This is unconventional and requires an explanation.

#### **Coercion and voluntariness**

Suppose that a patient faces the treatment options *A*, *B*, and *C*, that she ranks them in that order, but that she thinks of them as three good options, much unlike option *D*, *E*, and *F*, which she thinks of as bad options. Now, suppose also that the patient's doctor for some reason wants the patient to choose option *B*. Upon realizing that the patient is about to choose treatment *A*, the doctor bursts out: "If you choose *A*, I will reduce your dose of painkillers!" The patient, who fails to appreciate the peculiarity of the doctor's unprovoked outburst, shrugs disinterestedly. "Okay," she says, "I'll choose *B* then."

It is true that the patient chooses *B* non-autonomously, but although the doctor coerces her (per VaC standards), it is alien to how the patient leads her life to say that she chooses *B* non-voluntarily. Had the doctor forced the patient to choose between *D*, *E*, and *F*, it is likely that she would have thought of it as non-voluntary at the time of choosing, i.e., experienced an inner sense of resistance of the kind discussed above. But, since she thought of *A*, *B*, and *C* as good options, it did not dawn on her at the time that she might be wronged. If she later recollects the scene, she would be correct in saying that the doctor wronged her in some way, but incorrect in saying that she did not choose voluntarily; it would be an inaccurate description of something that should include her attitude toward the choice.

The example is intended to show that there is a difference between coercing someone who is engaged with a choice and coercing someone who is not. While it is possible, in most cases even likely, that a coerced agent is wronged in some way disregarding of whether she is engaged with her choice, the latter does not negate voluntariness. Whatever it is that makes it wrong to coerce people, the term "voluntariness" is not always the most suitable to employ in an articulation of the wrongness.

### Objections

The first objection targets the subjectivism of my proposal, the second is that my proposal is counterintuitive, and the final objection is that my proposal is unconventional. I will answer to them in that order.

There are two aspects of the objection that targets the subjectivism of my proposal; one is that subjectivism entails relativism, the other is that subjective experiences cannot be objectively evaluated. Whether an action is voluntary, it might be argued, should be determined by objective standards. Resting an account on individual perceptions is untrustworthy and allows for relativism. This objection is mistaken. First, the notion of thinking about an action as non-voluntary is not intended to single-handedly explain voluntariness, as all actions thought of as non-voluntary by the acting agent would then also be non-voluntary. It is one of two necessary and jointly sufficient criteria that in conjunction explains voluntariness (coercion being the other). Second, the fact that my notion includes *subjective judgments* does not entail that it builds on *subjectivism*. I have proposed an objective standard for what it is to think of an action as non-voluntary, and agents can fail to adhere to it. Thereby, my notion is not inherently untrustworthy and does certainly not allow for relativism.

Regarding the objection that subjective experiences cannot be evaluated, my answer is that a distinction must be made between *criteria of evaluation* and *the decision procedures* by which such criteria are evaluated. The criterion of evaluation is here the notion that an action is non-voluntary only if the agent thinks of it as such when being coerced to perform it. It rests on whether an agent experiences an inner sense of resistance, which is epistemically difficult to determine. This epistemic difficulty is here a matter of decision procedure, i.e., of deciding whether the criterion of evaluation is met in a particular instance. But, decision procedures are unrelated to the truth-value of criteria of evaluation. The truth-value of the notion that an action is non-voluntary only if the agent thinks of it as such when being coerced does not hinge on whether the agent's thoughts are epistemically available to an observer. The fact that my account of voluntariness is accompanied by epistemic difficulties does not mean that it is false.<sup>9</sup>

The next objection is that my proposal is counterintuitive. In most theories of voluntariness, coercion negates voluntariness by default. Nonetheless, I think that

---

<sup>9</sup> And, the epistemic *difficulty* does not necessarily amount to an epistemic *impossibility*.

there are strong reasons to adhere to the notion that an action is non-voluntary only if the agent thinks of it as such when being coerced. Therefore, regarding coercion, I will have to bite the bullet. In the above section, I have presented an example showing that it is perhaps not completely unreasonable that an agent can be subject to coercion and still act voluntarily. It is a matter of being (or not being) engaged with the choice in question; if the choice does not matter to the agent, coercion does not affect the voluntariness with which she chooses. Even if my arguments are unconvincing, it should be concluded that the standard view that coercion negates voluntariness by default should not be uncritically accepted.

The final objection is that my proposal is unconventional. Voluntariness is one of bioethicists' favorite go-to-concepts. However, I think that it is over-used. Other VaC theorists than those already mentioned have suggested that potentially constraining political, cultural, social, and economic factors may limit an agent's degree of control over her choices and thereby the voluntariness with which she chooses (Bull and Lindegger 2011, p. 27). Yet others have expanded the theory further. For instance, it has been shown that in research involving children many mothers consider decision-making by the child's father as socially normative, and some argue that this norm influences the mothers' voluntariness (Molyneux et al. 2005; Kamuya et al. 2011, p. 31). Lastly, in perhaps the most demanding version of the VaC theory, it has been argued that voluntariness requires authenticity, understood as choosing "in a way that is in conformity with one's identity, affective state, values, and goals, and is truthful to one's sense of self and view of the good life" (Berghmans 2011, p. 24). In other words, bioethicists have attempted to squeeze almost anything into the concept of voluntariness, as if it could explain all sorts of moral dilemmas.

However, the concept is not as applicable as bioethicists often seem to think. On the contrary, most actions are voluntarily performed, while potential moral problems associated with them are better explained through other concepts. Bioethicists should strive toward conceptual diversity and plurality, developing more concepts with better precision, rather than attempting to fit everything into one universal tool.

One tool that might have to be invented is a criterion of engagement, or interest, in informed consent. That is, if voluntariness should be understood more narrowly than what is common, which is my view, something else might need to fill the newly formed space in between already established informed consent criteria. Perhaps people cannot consent to a medical intervention or research participation if they are not engaged in the decision, in the sense discussed above.

**Putting the theory to work**

The question remains how to put my theory to work. I will briefly demonstrate that here.

The patient in the coffee example takes her medicine voluntarily but unknowingly. As she is unaware of the medicine in her coffee she does not consent to being medicated.

Had the nurse said, “drink your coffee or I will not let your family in during visiting hours tomorrow,” the patient would have been coerced. If the patient engages in the decision, as discussed above, and if the nurse had been successful in her effort to medicate the patient, it would have been an example of coercion that negates voluntariness. If the patient does not engage in the choice, but only shrugs disinterestedly, it would have been coercion that did not negate the voluntariness of her actions. Though, due to her ignorance of the medicine dissolved in the coffee, the patient’s actions would not amount to valid consent in either case.

Other examples are: A patient who hears voices in her head telling her to consent to treatment consents voluntarily but incompetently; a patient who consents to treatment because “everyone expects her to” consents voluntarily but non-autonomously; a patient who refuses treatment because she wants to die does so voluntarily.

**Concluding remarks**

This article is an attempt to stimulate bioethicist debate on voluntariness; its conditions, and role. It has been suggested that voluntariness should cover lesser moral ground than what is conventional. Voluntariness should be an agent-centered concept that takes the agent’s point of view into consideration through the notion that an act is non-voluntary only if the agent thinks of it as such when being coerced to perform it. Thereby, greater analytical precision can be achieved while voluntariness is reconnected with how people actually lead their lives, which fosters autonomy.

PERSONAL AUTONOMY AND INFORMED CONSENT

## Bibliography

- Ahlin, J. (forthcoming). "The impossibility of reliably determining the authenticity of desires: implications for informed consent." In: *Medicine, Health Care and Philosophy*. DOI: 10.1007/s11019-017-9783-0.
- Appelbaum, P. S., C. W. Lidz, and R. Klitzman (2009). "Voluntariness of consent to research: A conceptual model." In: *Hastings Center Report* 39.1, pp. 30–39.
- Archard, D. (2016). "The Methodology of Applied Philosophy." In: *A Companion to Applied Philosophy*. Ed. by K. Lippert-Rasmussen, K. Brownlee, and D. Coady. John Wiley & Sons, Ltd.
- Aristotle. (2014). *Nicomachean ethics*. Cambridge University Press.
- Banner, N. F. and G. Szmukler (2013). "'Radical Interpretation' and the Assessment of Decision-Making Capacity." In: *Journal of Applied Philosophy* 30.4, pp. 379–394.
- Beauchamp, T. L. (2010). "Autonomy and Consent." In: *The ethics of consent: Theory and practice*. Ed. by F. G. Miller and A. Wertheimer. Oxford: Oxford University Press, pp. 55–79.
- Beauchamp, T. L. and J. F. Childress (1983). *Principles of biomedical ethics*. 2nd ed. Oxford: Oxford University Press.
- (2013). *Principles of biomedical ethics*. 7th ed. Oxford: Oxford University Press.
- Berghmans, R. (2011). "Voluntary consent, normativity, and authenticity." In: *American Journal of Bioethics* 11.8, pp. 23–24.
- Björk, J., N. Lynøe, and N. Juth (2016). "Empirical and philosophical analysis of physicians' judgments of medical indications." In: *Clinical Ethics* 11.4, pp. 190–199.
- Bowman, D., J. Spicer, and R. Iqbal (2012). *Informed consent: A primer for clinical practice*. Cambridge: Cambridge University Press.
- Brennan, T. (1991). *Just doctoring: medical ethics in the liberal state*. University of California Press.
- Buchanan, A. E. and D. W. Brock (1990). *Deciding for Others: The Ethics of Surrogate Decision Making*. Cambridge University Press.
- Bull, S. and G. M. Lindegger (2011). "Ensuring consent to research is voluntary: How far do we need to go?" In: *The American Journal of Bioethics* 11.8, pp. 27–29.

- Burns, J. M. and R. H. Swerdlow (2003). "Right Orbitofrontal Tumor With Pedophilia Symptom and Constructional Apraxia Sign." In: *Archives of Neurology* 60.3, pp. 437–440.
- Charland, L. C. (2001). "Mental competence and value: The problem of normativity in the assessment of decision-making capacity." In: *Psychiatry, Psychology and Law* 8.2, pp. 135–145.
- (2015). *Decision-Making Capacity*. URL: <https://plato.stanford.edu/entries/decision-capacity/>.
- Christman, J. (2004). "Relational Autonomy, Liberal Individualism, and the Social Constitution of Selves." In: *Philosophical Studies* 117.1, pp. 143–164.
- (2009). *The Politics of Persons: Individual Autonomy and Socio-historical Selves*. Cambridge University Press.
- (2015). *Autonomy in moral and political philosophy*. URL: <http://plato.stanford.edu/archives/spr2015/entries/autonomy-moral/>.
- Colburn, B. (2008). "Debate: The concept of voluntariness." In: *Journal of Political Philosophy* 16.1, pp. 101–111.
- DeGrazia, D. (2005). *Human Identity and Bioethics*. Cambridge University Press.
- Dworkin, G. (1988). *The Theory and Practice of Autonomy*. Cambridge University Press.
- Elster, J. (1983). *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge University Press.
- Eyal, N. (2012). *Informed Consent*. URL: <http://plato.stanford.edu/entries/informed-consent/>.
- Faden, R. and T. Beauchamp (1986). *A History and Theory of Informed Consent*. Oxford University Press.
- Fisher, J. A. (2013). "Expanding the frame of 'voluntariness' in informed consent: Structural coercion and the power of social and economic context." In: *Kennedy Institute of Ethics Journal* 23.4, pp. 355–379.
- Frankfurt, H. (1971). "Freedom of the Will and the Concept of a Person." In: *The Journal of Philosophy* 68.1, pp. 5–20.
- Freedman, B. (1981). "Competence, Marginal and Otherwise: Concepts and Ethics." In: *International Journal of Law and Psychiatry* 4.1–2, pp. 53–72.
- Gillon, R. (2003). "Ethics needs principles—four can encompass the rest—and respect for autonomy should be "first among equals"." In: *Journal of Medical Ethics* 29.5, pp. 307–312.

#### BIBLIOGRAPHY

- Grisso, T., P. S. Appelbaum, and C. Hill-Fotouhi (1997). "The MacCAT-T: A Clinical Tool to Assess Patients' Capacities to Make Treatment Decisions." In: *Psychiatric Services* 48.11, pp. 1415–1419.
- Hippocrates. *Hippocratic oath*. URL: <http://medical-dictionary.thefreedictionary.com/Hippocratic+oath>.
- Ho, A. (2008). "The Individualist Model of Autonomy and the Challenge of Disability." In: *Bioethical Inquiry* 5.2–3, pp. 193–207.
- Hope, P. T. et al. (2011). "Anorexia Nervosa and the Language of Authenticity." In: *Hastings Center Report* 41.6, pp. 19–29.
- Hyman, J. (2013). "Voluntariness and choice." In: *The Philosophical Quarterly* 63.253, pp. 683–708.
- Jonsen, A. R. (2000). *A Short History of Medical Ethics*. Oxford University Press.
- Juth, N. (2005). *Genetic Information – Values and Rights. The Morality of Presymptomatic Genetic Testing*. Acta Universitatis Gothoburgensis.
- Kamuya, D., V. Marsh, and S. Molyneux (2011). "What we learned about voluntariness and consent: Incorporating 'background situations' and understanding into analyses." In: *The American Journal of Bioethics* 11.8, pp. 31–33.
- Kant, I. (1996). *Practical Philosophy (The Cambridge Edition of the Works of Immanuel Kant)*. Cambridge University Press.
- Katz, J. (1984). *The silent world of doctor and patient*. The Johns Hopkins University Press.
- Loss aversion. URL: <https://www.behavioraleconomics.com/mini-encyclopedia-of-be/loss-aversion/>.
- Lynøe, N. (1999). *Mellan cowboyetik och scoutmoral: Medicinsk forskningsetik i praktiken*. Liber.
- Meyers, D. T. (2001). "Authenticity for Real People." In: *Proceedings of the Twentieth World Congress of Philosophy: Philosophy of Mind and Philosophy of Psychology*. Ed. by B Elevitch. Philosophy Documentation Center, pp. 195–202.
- (2005). "Decentralizing Autonomy: Five Faces of Selfhood." In: *Autonomy and the Challenges to Liberalism*. Ed. by J Christman and J Anderson. Cambridge University Press, pp. 27–55.
- Millenson, M. L. (2011). "Spock, feminists, and the fight for participatory medicine: a history." In: *Journal of Participatory Medicine* 3.27.
- Miller, B. L. (1981). "Autonomy & the Refusal of Lifesaving Treatment." In: *The Hastings Center Report* 11.4, pp. 22–28.

- Miller, V. A. et al. (2011). "The decision-making control instrument to assess voluntary consent." In: *Medical Decision-making* 31.5, pp. 730–741.
- Molyneux, C. S. et al. (2005). "Even if they ask you to stand by a tree all day, you will have to do it (laughter)...!": Community voices on the notion and practice of informed consent for biomedical research in developing countries." In: *Social Science and Medicine* 61.2, pp. 443–454.
- Murray, P. M. (1990). "The History of Informed Consent." In: *The Iowa Orthopaedic Journal* 10, pp. 104–109.
- Nelson, R. M. et al. (2011). "The concept of voluntary consent." In: *The American Journal of Bioethics* 11.8, pp. 6–16.
- Nozick, R. (1969). "Coercion." In: *Philosophy, science, and method: Essays in honor of Ernest Nagel*. Ed. by W Morgenbesser. New York: St Martin's Press, pp. 440–472.
- (1974). *Anarchy, state, and utopia*. Blackwell Publishing.
- Olsaretti, S. (1998). "Freedom, force and choice: Against the rights-based definition of voluntariness." In: *The Journal of Political Philosophy* 6.1, pp. 53–78.
- (2008). "Debate: The concept of voluntariness—A reply." In: *The Journal of Political Philosophy* 16.1, pp. 112–121.
- O'Neill, O. (2003). "Some limits of informed consent." In: *Journal of Medical Ethics* 29, pp. 4–7.
- Oshana, M. A. L. (2015). "Is Social-Relational Autonomy a Plausible Ideal?" In: *Personal Autonomy and Social Oppression: Philosophical Perspectives*. Ed. by M.A.L Oshana. 1st ed. Routledge, pp. 3–24.
- O'Shea, T. (2011). "Consent in History, Theory and Practice." In: *Essex Autonomy Project Green Paper Report*. University of Essex: Essex Autonomy Project.
- Osuji, P. I. (forthcoming). "Relational autonomy in informed consent (RAIC) as an ethics of care approach to the concept of informed consent." In: *Medicine, Health Care and Philosophy*. DOI: 10.1007/s11019-017-9789-7.
- Siedentop, L. (2014). *Inventing the Individual: The Origins of Western Liberalism*. Penguin Books.
- Sjöstrand, M. and N. Juth (2014). "Authenticity and psychiatric disorder: does autonomy of personal preferences matter?" In: *Medicine, Health Care and Philosophy* 17.1, pp. 115–122.
- Strong, C. (2013). "Medicine and Philosophy: The Coming Together of an Odd Couple." In: *The Development of Bioethics in the United States*. Ed. by J. R Garret, F Jotterand, and D Ralston. Springer Netherlands.

#### BIBLIOGRAPHY

- Sutrop, M. (2011). "Changing Ethical Frameworks: From Individual Rights to the Common Good?" In: *Cambridge Quarterly of Healthcare Ethics* 20.4, pp. 533–545.
- Tan, D. J. O. A. et al. (2006a). "Competence to make treatment decisions in anorexia nervosa: thinking processes and values." In: *Philosophy, Psychiatry, & Psychology: PPP* 13.4, pp. 267–282.
- (2006b). *Competence to make treatment decisions in anorexia nervosa: thinking processes and values*. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2121578/pdf/nihms5623.pdf>.
- Tännsjö, T. (1999). *Coercive Care: The Ethics of Choice in Health and Medicine*. Harvard University Press.
- Taylor, J. S. (2009). *Practical Autonomy and Bioethics*. Routledge.
- Velleman, D. J. (2002). "Identification and Identity." In: *Contours of Agency: Essays on Themes from Harry Frankfurt*. Ed. by S Buss and L Overton. The MIT Press, pp. 91–123.
- Wall, E. (2001). "Voluntary action." In: *Philosophia* 28.1–4, pp. 127–136.
- Wertheimer, A. (2012). "Voluntary consent: Why a value-neutral concept won't work." In: *Journal of Medicine and Philosophy* 37.3, pp. 226–254.
- Widdows, H. (2011). "Localized past, globalized future: towards an effective bioethical framework using examples from population genetics and medical tourism." In: *Bioethics* 25.2, pp. 83–91.
- Wiking, E. (2012). *Health Care on Equal Terms for Immigrants in Sweden*. Karolinska Institutet.
- Winnicott, D. W. (2007). "Ego Distortion in Terms of True and False Self." In: *The Maturation Process and the Facilitating Environment: Studies in the Theory of Emotional Development*. Karnac Books, pp. 140–154.
- Wolff, J. (2011). *Ethics and Public Policy: A Philosophical Inquiry*. Routledge.