# 14 Does Closed-Loop DBS for Treatment of Psychiatric Disorders Raise Salient Authenticity Concerns?

Ishan Dasgupta, Andreas Schönau, Timothy Brown,
Eran Klein, and Sara Goering

## 14.1 Introduction

Deep brain stimulation (DBS) is used by over 160,000 people worldwide to treat a range of neurological disorders including Parkinson's disease (PD), essential tremor, and dystonia [1]. By implanting electrodes in specific brain regions, clinicians are able to provide stimulation in order to manage symptoms that are not responsive to traditional pharmacological approaches. Despite these successes, thus far DBS technologies have shown limited efficacy for treating psychiatric disorders such as depression, eating disorders, and addiction [1]. This has sparked immense interest in developing the next generation of DBS technologies to treat psychiatric disorders. Where the current DBS systems are *open-loop*—applying a constant or intermittent electrical current to the brain—the new generation of DBS devices will utilize artificial intelligence technologies, such as machine learning algorithms, to facilitate *closed-loop* implants that are adaptive and continuously modified by neural feedback [2]. Closed-loop DBS devices read neural data, interpret the signals to make a clinical decision, and stimulate the brain dynamically. This process occurs continuously and without active input from users or clinicians, allowing for far fewer adjustments and improving treatment specificity.

I. Dasgupta (✉) · A. Schönau · T. Brown · S. Goering
Department of Philosophy, University of Washington, Seattle, WA, USA

Center for Neurotechnology, University of Washington, Seattle, WA, USA
e-mail: idasg@uw.edu

E. Klein
Department of Philosophy, University of Washington, Seattle, WA, USA

Center for Neurotechnology, University of Washington, Seattle, WA, USA

Department of Neurology, Oregon Health Sciences University, Portland, OR, USA



F. Jotterand, M. Ienca (eds.), *Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical & Policy Issues*, Advances in Neuroethics,
https://doi.org/10.1007/978-3-030-74188-4_14

Since we invite the possibility of automating the modulation of our brain, we must also consider the ways in which such devices may change the end user and the way they interact with their world. These worries are backed up by empirical evidence showing that at least some users of open-loop DBS report changes to aspects of their personal identity, agency, or sense of self [3–6]. These reports have been the catalyst for a robust and still ongoing debate over implanted neurostimulators and the nature of the impact they have on users [7]. For example, is it worth treating an illness if people no longer feel like themselves? What if friends and loved ones no longer recognize the person they knew? Some of these worries are not entirely new—they have been raised around the issue of authenticity and the use of antidepressants [8] or open-loop DBS devices [9]. On the other hand, some have noted that closed-loop DBS technologies raise *salient* authenticity concerns, especially when they are used to treat psychiatric conditions [10, 11].

In this chapter, we explore the ways in which closed-loop DBS systems can introduce changes to the self that are different from open-loop DBS. At each step of the closed-loop process—reading neural signals, interpreting data, and stimulating the brain—new complexities around authenticity are introduced by the increased reliance on automated systems. Threats to privacy, traditional clinical relationships, and agency raise concerns about whether users can still live authentically. These concerns can be mitigated by developing new ethical guidance to address the unique setting of closed-loop DBS.

## 14.2 Deep Brain Stimulation and Psychiatric Disorders

The treatment of psychiatric disorders has taken many forms over history. Starting with psychoanalytic psychotherapy, the field of psychiatry eventually developed pharmacological and surgical interventions. The emergence of DBS technologies in the early 1990s [12, 13] introduced an alternative to ablative psychosurgeries for patients with a wide range of refractory psychiatric disorders. DBS offered the ability to deliver electric stimulation "into specific targets within the brain and the delivery of constant or intermittent electricity from an implanted battery source" [1]. Over the last 30 years, the use of DBS has been investigated in the treatment of psychiatric conditions such as major depressive disorder (MDD), obsessive compulsive disorder (OCD), bipolar disorder, Tourette syndrome, schizophrenia, addiction, and anorexia [1].

Although the exact mechanism differs according to the disease being treated, the basic concept behind open-loop DBS for psychiatric disorders is similar. Researchers identify a candidate for surgery based on a symptom-based tool like the DSM V, a cognitive measure, such as the Montgomery-Asberg Depression Rating Scale (MADRS), and history of inadequate response to other forms of treatment. In the well-known RECLAIM trial for MDD, participants must have had failed treatment trials with at least four antidepressants [14]. If the patient meets the criteria for a clinical trial, electrodes (one or two, depending on the study) are placed in a specific region of the brain thought to be responsive to stimulation. After the surgery, stimulation is turned on and the voltage set in order to manage the symptoms of the condition. Stimulation is adjusted in trial visits based on assessments of symptoms and side effects. This stimulation is either continuous or follows a predetermined schedule. In certain cases, users may have individual control over their stimulation and may be able to turn the device on and off, as well modulate the voltage setting.

DBS offers large improvements over prior psychosurgeries, such as being able to directly interface with brain circuitry while avoiding permanent lesioning. Due to its efficacy, DBS, as opposed to ablative psychosurgery, is now considered the treatment of choice for some individuals with treatment refractory PD, essential tremor, and epilepsy [15]. Regrettably, there has been limited success in using DBS to treat psychiatric disorders. Initial studies using DBS for OCD and MDD showed promise, but randomized clinical trials failed to show similar efficacy [14, 16]. Researchers argue there are explanations for the discrepancy in results [2, 17, 18]. Our understanding of neural circuitry responsible for mental illness is lacking when compared to motor disorders. Additionally, psychiatric disorders are often the result of multiple dysfunctional circuits as opposed to one network. For instance, what we classify as MDD may actually be a collection of unique conditions. Furthermore, the lack of reliable biomarkers for symptoms makes it difficult to determine whether modulation is responsible for clinical success. Taken together, researchers argue that more targeted stimulation that is modulated according to neural feedback across different neural circuits may be a solution to using DBS to treat psychiatric disorders.

Closed-loop DBS systems have been proposed as an alternative to open-loop DBS and can succeed where their predecessors failed [17]. While open-loop DBS works in a unidirectional fashion by providing a constant level of stimulation to targeted brain areas without integrating any sort of neural feedback, the new generation of DBS devices is "closing the loop" by recording neural data, analyzing it for salient features utilizing machine learning algorithms, and using these analyses to alter stimulation parameters like amplitude and frequency. This loop of reading data, analyzing data, and stimulation addresses some concerns about authenticity raised by open-loop DBS but raises salient issues in the closed-loop context.

## 14.3 Authenticity and Treatment of Psychiatric Disorders

In a colloquial sense, to be authentic means to be "true to oneself." For many, being true to oneself is not a complicated calculus. We all can intuitively point to something about ourselves that is more foundational than everything else. For example, one can change one's hair color easily, but it is much more difficult to become an outgoing person if one is born as an introvert. On the other hand, we also understand that people change their conceptions of self over time and are not permanently bound to traits they acquired early in life. A once lazy student who becomes more disciplined about coursework by using a new mindfulness technique can reinvent her identity to become a hardworking person. As long as she is able to conform her actions to her new conception of self as a responsible student, it seems uncontroversial that most people in her life would be willing to grant that she has changed a part of her former identity. This example illustrates that choices affecting authenticity run a spectrum from minor to fundamental. It also shows that there are different ways of thinking about authenticity. Since a comprehensive discussion of all types of authenticity are beyond the scope of this chapter, below we highlight three ways that have been prominently discussed in the neuroethical literature on DBS.

### 14.3.1 Sense of Authenticity

As early as 2006, Schupbach et al. reported that patients receiving deep brain stimulation (DBS) for PD felt alienated after receiving their device even though they experienced a reduction in measured clinical symptoms: “Now I feel like a machine, I’ve lost my passion. I don’t recognize myself anymore” [5]. Recent studies of DBS for psychiatric indications have echoed similar concerns. A study by De Haan et al. discusses one DBS user with OCD who felt like a different person because of the change in their libido following implantation: “I did not like that at all…No, that clearly didn’t fit with who I am…it was really too much; that really wasn’t me, you know; I really felt as if there was someone [else] standing next to me…” [11]. Goering et al. interviewed end users, some of whom expressed that they were no longer able to act in ways that were consistent with their pre-DBS self: “I can’t really tell the difference. There are three things—there’s me, as I was, or think I was; and there’s the depression, and then there’s depression AND the device and, it, it blurs to the point where I’m not sure, frankly, who I am” [10, 11]. This empirical data suggests, at the very least, that DBS can cause changes to users that make them feel as if they are no longer like their prior self.

One way we think about authenticity is in terms of our sense of self *in a particular moment*. Does changing the color of my hair make me feel more like my authentic self or like an imposter? This sense of authenticity captures the way in which an individual experiences their self from the first-person perspective. A sense of authenticity is paramount to ensuring that people feel comfortable living their everyday lives. It is also the type that is most intuitively thought of when discussing DBS technologies. For example, after a patient undergoes implantation, how they feel once stimulation begins is of significant concern. DBS devices differ from pharmacological interventions in that the effects are typically sudden and can be jarring in nature rather than slow and iterative [13]. It is possible that feelings of alienation are due to the lack of time users have to adjust to major changes in mood. Many traditional pharmacological treatments, such as SSRIs, take weeks to months to reach their full effects allowing users to experience changes more gradually [19]. In this context, it may be difficult to determine whether feelings of inauthenticity are caused by the stimulation from the DBS itself or from changes the user experiences to their personality due to amelioration of a disease condition. That is to ask, is the device stimulating areas of the brain in ways that cause these new changes, or are these changes the result of reducing the negative effects of depression? [11]. Regardless of the cause of the changes, there is a very real sense in which end users of open-loop DBS experience feelings of inauthenticity.

### 14.3.2 Narrative Authenticity

A second way of thinking about authenticity is how the person aligns their actions with their true self [20, 21]. On a practical level, some of the worries about end users' feelings of inauthenticity after receiving DBS are tied to the prevalence of the concept of autonomy in modern bioethics. Generally, the worry is that a person will be autonomous, capable of understanding the consequences of their actions and choosing freely, but not authentic [22]. For example, in a classic example from the literature, a man undergoing DBS for PD becomes manic and develops a new interest in gambling [23]. When the device is turned on, the man claims that this new version of himself, where he is more impulsive and risk taking, is his real self and that his former self was an inauthentic version. Thus, if our gambling man is deemed to have retained his autonomy, under a traditional bioethical inquiry it may be inappropriate to deny him the right to make all medical decisions on his behalf. On the other hand, giving autonomy more weight than authenticity may presume an inappropriate dichotomy between the two. Under this view, it may be impossible for one to act autonomously if they cannot act in a way that is consistent with their true self. If we are to subscribe to this view, however, in what ways would we integrate authenticity into our current legal system for informed consent and medical decision-making? For example, at present, we have an established framework for determining whether a person is capable of autonomous behavior and can be legally responsible for their own medical decision-making [24]. Integrating authenticity would require that in cases like the man who gambles, clinicians would have some power to restrict the man's decision-making if his actions are not consistent with his prior self.

### 14.3.3 Assessing Authenticity

A third way of thinking about authenticity is how to determine whether someone is truly acting authentically. The way one approaches this question has to do with their views regarding the nature of the true self. Those who hold an essentialist view of the self argue that the self is composed of an unchanging and fixed core that represents one's true self [25]. Any departures from this fixed self would result in an inauthentic self. Those with existentialist views generally believe that the self is ever-changing and is constantly redefined by the individual through self-determination [25]. Under this view, people are free to change their true self throughout their life and can reinvent themselves if they choose. Finally, according to relational conceptions of the self, identity is constructed through repeated recalibration based on feedback from others, rather than in isolation [26]. These views are not mutually exclusive, and most people will likely subscribe to some hybrid version of the self that incorporates aspects from each framework. We do not attempt to settle this long-standing debate here. Rather, we argue that irrespective of which view one holds, in practice there are often disagreements between the user, their family members, and clinicians about whether or not the person is acting authentically after receiving treatment for psychiatric disorders.

Although there is an understandable preference to defer to the end user in these cases, there may be practical value in involving others in clinical decision-making when determining authenticity. Feedback from friends and family may help a clinician determine whether an end user will effectively integrate a device into their life. Given that most patients receiving these devices are likely to continue to have caregivers substantially involved in their daily lives, if DBS creates changes in personality that are too extreme, it may not be advisable to go through with the procedure even if it offers reduction of clinical symptoms. That being said, it is worth asking whether we are only concerned about feelings of inauthenticity when third-party perspectives disagree with one's first-person narrative? If the gambling man had become more thoughtful and loving, would clinicians and family members object? These are some of the questions that arise when considering issues related to assessing authenticity.

As discussed above, one way to evaluate feelings of inauthenticity is to rely on first-person reports from end users. Whether one believes the true self is fixed or changing, the end user will ultimately be in the best position to determine what is her true self and in what ways she wants to change that self and embrace a future identity. For example, if the gambling man is content with his new identity and is not breaking laws or harming himself or others, some may argue that he has a right to reinvent his new identity (even if those around him would prefer that he not). Furthermore, if this change is coupled with an improvement in clinical outcomes, one might argue that the change in the man's personality is a risk worth taking given the benefit he is experiencing from the DBS. On the other hand, what is to be made of the man's values and desires before he received the implant? For example, what are we to do if upon turning the DBS off we find that the man reflects upon and rejects his recent impulsive self? We are then left with two first-person accounts—both of which feel like they are authentic. Along these lines, a difficulty has emerged in the neuroethics literature about how to objectively determine when an end user's self-reported claims of inauthenticity are to be trusted [27]. How are clinicians and family members to resolve the discrepancy between the calm and thoughtful person they knew before implantation and the rash and offensive man they see before them now?

### 14.3.4 Why Authenticity?

Before starting our discussion of the differences between open and closed-loop DBS technologies, we must briefly consider why authenticity, as opposed to other ethical principles, should be the focus of our inquiry. We focus on authenticity for a few reasons. First, the empirical evidence, as we have seen above, illustrates that when end users talk about changes to their self, many intuitively utilize the language of authenticity. Second, authenticity is often a central concern expressed by people suffering from psychiatric disorders [28]. Psychiatric disorder can often lead to people feeling less like themselves, and treatment with DBS is sought in order to improve authenticity. In this way, when we consider closed-loop systems for treatment, we must examine both how they might reduce symptoms of the disease that implicate authenticity and the ways in which the device itself introduces new concerns about authenticity. Third, authenticity can offer practical utility in guiding clinical decision-making [29]. Since traditional bioethical inquiry heavily favors concepts such as autonomy as being paramount for medical decision-making, authenticity may offer a way to check our typical intuitions about treatment decisions. Finally, using DBS for psychiatric disorders, rather than motor or sensory disorders, raises an additional layer of complexity because treatment will seek to change the users' beliefs, desires, and emotions—key components of their self—directly rather than such changes occurring as an unintended consequence [11]. Therefore, authenticity, rather than autonomy, or other bioethical principles, may help us better understand the journeys end users take in coming to terms with the changes to the self-introduced by DBS.

## 14.4 Authenticity and Closed-Loop DBS

In what follows, we discuss three different technological improvements made by closed-loop DBS and how they may alleviate or exacerbate the types of authenticity we discussed earlier.

### 14.4.1 Reading Data

Unlike other treatments used for psychiatric disorders, closed-loop systems will directly measure neural data from the surface of the brain as opposed to measuring symptoms based on clinical observation and anecdotal reports [2]. This introduces a new range of possible data that is considered when treating psychiatric disorders. Traditionally, clinicians are limited to subjective evaluations and indirect behavioral assessments to monitor clinical progress. For example, when a patient with MDD is prescribed a new SSRI, a clinician typically waits for some time before conducting a clinical assessment of the drug's efficacy. Similarly, in open-loop DBS, the clinician is using clinical data to adjust the stimulation parameters in order to get the desired reduction in symptoms. The process, in both pharmacological and open-loop DBS, involves much trial and error and can be burdensome for end user and clinician alike [30].

In contrast, in closed-loop DBS, the device will measure neural activity related to a disease symptom directly from the brain. For instance, neural signals gathered by implantable ECoG electrodes can potentially be used to measure dysfunctional states in neural networks associated with psychiatric disorders [2]. The goal would be to develop a list of common phenotypes that are associated with the disorder and then try to identify these phenotypes autonomously. The ability for closed-loop devices to read neural data in this way can help to improve the treatment of psychiatric disorders and thereby reduce issues related to authenticity that are the result of the disease.

First, since the device is continuously reading out brain data, the gathered information is valuable for providing ongoing health monitoring and screening. This real-time characteristic of the system may be used to offer preventive care. Here, individual brain readings are potentially useful for early detection and diagnosis of other related diseases that may be leading to feelings of inauthenticity in the end user. For instance, if a patient is implanted with a closed-loop DBS to treat depression, the monitoring of neural data may alert the clinician to a comorbid condition that needs treatment. An early intervention can help the patient avoid further feelings of inauthenticity as a result of their disease. The additional neural data can also be used for reconfiguring settings (see below) that make the patient feel less alienated, rather than having to wait for an appointment where one has to describe the ways in which they feel different. Second, on a large scale, aggregation of neural data on a population level can accelerate our understanding of neuroscience by harnessing big-data and machine learning algorithms [18]. This gathered data might be used to understand the underlying biology of psychiatric disorders [18]. In the long run, this cumulative knowledge from numerous patient cohorts might help to generate a better conceptual understanding of the brain functions that implicate feelings of inauthenticity. This could improve treatment by tailoring it towards the end user's reports of how they feel.

However, one current obstacle, especially for psychiatric disorders, is the lack of reliable biomarkers for usefully differentiating between pathological and healthy states [18]. Closed-loop devices will only be as useful if they are able to measure data that is relevant to a user's condition. If these devices are unable to measure accurate neural phenotypes, they may be as (in)effective as our current observation-based and self-reported measures of mental illness.

Another issue raised in the literature when it comes to ongoing recordings is the exploitation of sensitive data. In order to process the large amounts of data collected and for clinicians to be able to make remote adjustments, closed-loop devices may have to be connected wirelessly to external systems that have more computing power or allow clinician access. This may make them vulnerable to hackers, who can potentially access private information and even control the device [31]. In the literature, this risk is discussed under the umbrella term of "neurocrime" that refers to individuals aiming at "illicit access to and manipulation of neural information and computation." [32]. Here, hackers might gain access to information that represents the patient in a way he feels uncomfortable in being represented with. Exploiting this information can change interpersonal dynamics of privacy that may result in authenticity issues. If we take seriously the likelihood that closed-loop devices will introduce new security issues, hackers could not only steal private information but also alter stimulation (see below) in a way that affects motor function, impulse control, and emotions [33]. Here, the ability to externally change stimulation patterns can be exploited to intentionally impact the end user's experience of authenticity. Designing future medical devices with a specific standard of neurosecurity that implements relevant security principles could help to provide neural devices with adequate security mechanism [34].

### 14.4.2 Analyzing Data

Once the patient's data is read, it will be analyzed through an algorithm that utilizes artificial intelligence or machine learning technologies. The closed-loop device must correctly interpret the user's neural data and associate it with a corresponding disease state. If the appropriate criteria are met, the system will determine that the user is experiencing symptoms of their illness and that stimulation needs to be provided. In the pharmacological and open-loop context, a clinician would utilize the DSM to tabulate a patient's anecdotes in order to make a diagnosis or would rely on observation.

There are two important implications for this stage of the process. First, making disease assessments using various biomarkers and data sources instantaneously allows for the system to make individualized diagnosis based on a user's unique brain activity. Second, population level data will be utilized to train the system so that it recognizes brain activity that is considered dysfunctional as indicated by it falling outside the normal range found in the training set. Both implications mark a considerable step forward compared to the manual tinkering necessary in open-loop devices. Patients report that the trial-and-error phase for adjusting the stimulation to meet their individual needs is one of the most frustrating aspects during clinician visits [3].

Closed-loop devices allow us to directly tackle this issue. Here, the device takes the role of the clinician in changing parameter settings but does so infinitely more times than a person could. In theory, the patient can benefit from less personally demanding visits for recurring parameter adjustments in clinical settings [35]. At the same time, the adaptivity of the system may result in more effective, personalized treatments that prevent potential overstimulation and unintended side effects [18]. For instance, in the case of DBS for PD, stimulation can be turned off when the patient is at rest while it can be turned back on when a new movement is recognized. Here, tailoring the system towards the current needs of the individual through adaptive data analysis is a tremendous improvement from the ongoing stimulation provided by open-loop DBS. Since the system permits the user to forget about the ongoing stimulation, the discontinued need to constantly reflect about their current state might prove helpful to provide a background setting in which the user feels comfortable, presumably resulting in a more authentic state.

On the other hand, there are ways in which we can imagine a system that makes diagnoses based on machine learning algorithms may introduce bias against certain types of people. For example, what data-set will the artificially intelligent systems be trained on? There are ethical concerns about how our existing societal biases might be integrated into machine learning algorithms and become further automated, making them more difficult to recognize or address [36]. These biases are not limited to gender or race and can include negative views regarding people who do not exhibit neurotypical behavior [37]. There are ways in which closed-loop systems can undermine a user's authenticity if they misdiagnose legitimate feelings or emotions as being pathological rather than justified in a particular situation. For example, what if a user is angered as a result of experiencing racial injustice. Would a closed-loop device trained on users who never experienced that type of stress be able to distinguish between anger felt in response to a legitimate trigger from anger that is due to dysfunction in neural networks? These types of misdiagnoses have the potential to make some users feel as if the system is forcing them into behavior that is inauthentic.

### 14.4.3 Stimulation

Adaptive systems will utilize the results of data analysis to formulate a treatment plan. This will be based on two levels of data: personal data indicating an individual's normal variation in symptoms and population level data indicating the range of functional neural activity. In this way, the device can calculate the stimulation needs of an individual user within the proven range for the therapy. This will allow the closed-loop system to provide a more precise stimulation at multiple brain locations at varying voltages without the need for clinical manipulation. As in prior sections, the promise of closed-loop devices is that they may mitigate concerns about authenticity by offering better treatment of underlying psychiatric disorders.

Executing this treatment plan, however, relies on the human brain to simultaneously adapt to ongoing changes in stimulation. This immediate connection between end user and device can lead to problems of authenticity that may be manifested in feelings of alienation [6] or autonomy concerns [22]. From a phenomenological perspective, some patients report an impact on their perception of themselves and their bodies, resulting in statements that they are feeling like a robot or like an electric doll [5]. Here, the pressing neuroethical issue consists in the difficulty a user has in differentiating between what he is doing and what the device is doing.

In the literature, this concern is addressed as the potential danger of closed-loop DBS to undermine agency in a way that the individual's capacity to regulate mood may not allow the normal range of emotional responsiveness [10, 11]. Imagine a patient with a psychiatric DBS who attends a funeral but is not able to produce the expected emotion of crying. Since the stimulation is delivered automatically, the patient would not be able to control the device or change any settings intentionally. Even worse, the patient might not even be able to recognize whether the current emotional setting is a consequence of the device's adaptive stimulation or actually a part of his "original" self without the device. As a result, the device being always on with its adaptive, self-learning mechanisms may put the user into a state of constant uncertainty. While a user can always stop taking prescribed pills when problems occur, an equivalent action of stopping the stimulation may not be possible while using closed-loop DBS [10, 11]. While a drug cannot change its method of modification mid-treatment to account for unwanted side effects, a closed-loop device could make it so the user never even feels alienated. The AI-assisted device, so it seems, is always in control since it autonomously determines the time and intensity of the stimulation. Even when it is not stimulating, it is doing so for a reason that is unbeknownst to the user. This shows that, especially in psychiatric DBS, the implementation of an AI is not necessarily beneficial for the individual's experience of living with the device. Instead, the addition of another layer of treatment decisions utilizing an integrated AI potentially exacerbates the user's uncertainty about who is responsible for mainting their well-being.

This uncertainty may also motivate concerns about whether the user is in control over their actions. Imagine a closed-loop DBS causing a cramp by overstimulating a patient who, as a consequence, unintentionally turns the steering wheel of his car. Who is responsible for the resulting accident? The patient, clinician, or the programmer of the AI? This general issue of responsibility ascription in stimulating neural devices is coined in the literature as a "responsibility gap" [38]. The integration of AI into this control scheme adds additional complexity to the already existing black box nature of neural devices, which makes recognizing and understanding the inner workings of the closed-loop device and the influence it has on the patient immensely difficult. For our current discussion, we are less concerned about who is morally or legally responsible for an involuntary action caused by a closed-loop DBS, but rather how the user may *feel* less in control.

There are two ways in which the user may feel less control. First, they may feel as if the device is acting instead of them leading to feelings of alienation. Second, if one feels less in control, they may also be less motivated to enact positive behavioral change in their lives thereby implicating issues related to narrative authenticity. For example, it is possible that users, even if they feel that they are in control over their actions, may nonetheless cede responsibility to the device because they believe that an artificially intelligent system will address all their treatment needs. A patient suffering from depression who may have exercised in the past to improve their mood in conjunction with an SSRI may now feel less motivated, since they have a system that provides more precise stimulation and obviates the need for the user to take personal responsibility for improving their mood. This raises concerns about whether someone is still authentic when they no longer have to put in the hard work of improving themselves [8].

In this case, one might note that the patient can always check in with the physician if there are problems with the stimulation parameters. As a response, the clinician might check whether the device is stimulating correctly according to its readings, but will be unable to significantly change the programming, since that will probably require updates or changes made by the manufacturer. Furthermore, since the device is stimulating solely on the basis of neural data that does not take the patient's unique experience of having a psychiatric disorder into account for providing treatment, it is unlikely to be able to handle the task of processing the patient's phenomenological experience when calculating stimulus [17]. Here, clinicians need to be sensitive about keeping a personal relationship with their patients by including their anecdotally related symptoms into the therapeutic process [35].

## 14.5 Future Directions

We have identified several features of closed-loop DBS that may impact users' feelings of authenticity in morally salient ways. Our discussion elucidated that closed-loop devices, if developed to their full potential, can reduce concerns about authenticity by reducing disease-related symptoms. In this way, users may feel more like themselves, be able to better construct their own narrative identities because of improved mood or control over their behavior, and others may start to recognize the person they were before being impacted by disease. Conversely, this discussion also highlighted the ways in which new technological aspects of closed-loop systems raise salient concerns about authenticity. Reading of neural data can raise issues about whether the correct information is being recorded and opens the door for unauthorized access to neural data. The machine learning systems involved in analyzing the data can introduce a variety of biases into the systems, thereby automating many of the problems we already have in treating and diagnosing mental illness. Finally, adaptive stimulation can exacerbate existing worries about DBS causing changes in the self, as well as introducing new worries about whether a user will be in control over their own actions and future behavior.

To ensure safe and responsible use of this emerging technology, it is essential that we not only predict these future issues, but also flesh out initial steps towards possible solutions. Here, a potential first step would be to create ethical guidelines that structure the ongoing debate in a way that any anticipated negative consequences of closed-loop technology are prevented or mitigated significantly. One way forward could be to recognize that authenticity has an important function in medical decision-making without giving it more power than it currently holds. Since our legal systems are not yet equipped to deal with issues related to authenticity, we can attempt to deal with feelings of inauthenticity through additional care for end users while keeping autonomy as the dispositive concept when it comes to deciding whether a person is able to make their own decisions.

Second, since closed-loop systems make use of algorithms that require training and including users in the training process can either foster authenticity or exacerbate inauthenticity, potential end users should get a thorough explanation of what will (or might) happen to them and what they can do if they feel alienated, isolated, or odd post-surgically. In order to achieve this, clinicians should make sure that their patients are well-informed about potentially occurring changes in authenticity by offering background knowledge about all three stages (reading, analyzing, and stimulation) of the neural device. Additionally, ethical guidance should focus on the ways in which developers of closed-loop systems are cognizant of the different biases that can be introduced into their devices if careful attention is not paid to the data sets they are trained on. Guidelines should be developed in order to guide developers when they are creating these devices so that they include enough training data in order to capture a multitude of diverse neural data.

Third, as touched on above, the relationship between the patient and clinician will need adapting as closed-loop systems become introduced in clinical practice [38]. Here, patients should be given the option of participating in the tuning process if they desire it. This may include deeper training on how their stimulator works, how its algorithms work, what data is collected, how the data is analyzed, and how the stimulation is influencing different parts of the brain. In terms of privacy, this may include an overview on the measures that are in place to protect their recorded brain data and secure their implant from security breaches. In terms of alienation, this could include thoroughly informing the patient about possible changes on a psychological level as well as providing guidance on who end users can turn to if they experience a sudden change in authenticity.

Another way clinicians can help users guard against potential unwanted changes to the self is to utilize legal instruments like advanced directives. For example, Klein has argued that in cases like the gambling man, the user may use a Ulysses contract to prospectively note behaviors they find unacceptable and situations where they would want their device turned off, even if this goes against the will of the post-DBS person: "If I ever become a compulsive gambler, please intervene" [39]. Named after the hero from Homer's Odyssey who famously instructed his crew to leave him tied to the mast and ignore his future self, Ulysses contracts could serve as a DBS analog to advanced directives used in patients with dementia. Utilizing Ulysses contracts in conjunction with preimplantation consultation with family members and loved ones could help the user establish a series of conditions under which clinicians would remove the treatment against the patient's wishes. This could include the ability for people close to the user to raise a red flag if they sense the person's behavior is changing but the user themselves does not notice the change or does not mind the change. This can guard against concerns about assessing authenticity and determine which autonomous self to respect when there is a conflict between a person before and after they receive a closed-loop DBS.

Finally, we saw earlier that authenticity, along with identity more broadly, is a relational concept; people evaluate the authenticity of their actions with the help of others, and what counts as an authentic action is constrained by others as well. Closed-loop DBS users could look to other users to figure out if their actions are authentic to them, or if they are a by-product of how the device operates. Research on and development of closed-loop devices should facilitate communication between users of these devices by coordinating meetups and recording honest testimonials (not just in the form of positive marketing materials).

While these and other considerations still need to be discussed more thoroughly in the ongoing debate, once fleshed out and put into place, they offer valuable support for end users to successfully adapt to living with closed-loop devices.

## References

1. Lozano AM, Lipsman N, Bergman H, Brown P, Chabardes S, Chang JW, et al. Deep brain stimulation: current challenges and future directions. Nat Rev Neurol. 2019;15:148–60.
2. Widge AS, Malone DA, Dougherty DD. Closing the loop on deep brain stimulation for treatment-resistant depression. Front Neurosci. 2018;12:175.
3. Klein E, Goering S, Gagne J, Shea CV, Franklin R, Zorowitz S, et al. Brain-computer interface-based control of closed-loop brain stimulation: attitudes and ethical considerations. Brain Comput Interfaces. 2016;3:140–8.
4. de Haan S, Rietveld E, Stokhof M, Denys D. Effects of deep brain stimulation on the lived experience of obsessive-compulsive disorder patients. PLoS One. 2015;10(8):e0135524. [cited 2015 Dec 10]; http://philpapers.org/rec/DEHEOD.

5. Schüpbach M, Gargiulo M, Welter ML, Mallet L, Behar C, Houeto JL, et al. Neurosurgery in Parkinson disease a distressed mind in a repaired body? Neurology. 2006;66:1811–6.
6. Kraemer F. Me, myself and my brain implant: deep brain stimulation raises questions of personal authenticity and alienation. Neuroethics. 2013;6:483–97.
7. Mackenzie C, Walker M. Neurotechnologies, personal identity, and the ethics of authenticity. In: Clausen J, Levy N, editors. Handbook of neuroethics [Internet]. Dordrecht: Springer; 2015. p. 373–92. https://doi.org/10.1007/978-94-007-4707-4_10.
8. Elliott C. Enhancement technologies and the modern self. J Med Philos. 2011;36:364–74.
9. Pugh J, Maslen H, Savulescu J. Deep brain stimulation, authenticity and value. Camb Q Healthc Ethics. 2017;26:640–57.
10. Goering S, Klein E, Dougherty DD, Widge AS. Staying in the loop: relational agency and identity in next-generation DBS for psychiatry. AJOB Neurosci. 2017;8:59–70.
11. de Haan S, Rietveld E, Stokhof M, Denys D. Becoming more oneself? Changes in personality following DBS treatment for psychiatric disorders: experiences of OCD patients and general considerations. PLoS One. 2017;12:e0175748.
12. Gardner J. A history of deep brain stimulation: technological innovation and the role of clinical assessment tools. Soc Stud Sci. 2013;43:707–28.
13. Schechtman M. Philosophical reflections on narrative and deep brain stimulation. J Clin Ethics. 2009;21:133–9.
14. Dougherty DD, Rezai AR, Carpenter LL, Howland RH, Bhati MT, O'Reardon JP, et al. A randomized sham-controlled trial of deep brain stimulation of the ventral capsule/ventral striatum for chronic treatment-resistant depression. Biol Psychiatry. 2015;78:240–8.
15. Meidahl AC, Tinkhauser G, Herz DM, Cagnan H, Debarros J, Brown P. Adaptive deep brain stimulation for movement disorders: the long road to clinical therapy. Mov Disord. 2017;32:810–9.
16. Malone DA, Dougherty DD, Rezai AR, Carpenter LL, Friehs GM, Eskandar EN, et al. Deep brain stimulation of the ventral capsule/ventral striatum for treatment-resistant depression. Biol Psychiatry. 2009;65:267–75.
17. Widge AS, Ellard KK, Paulk AC, Basu I, Yousefi A, Zorowitz S, et al. Treating refractory mental illness with closed-loop brain stimulation: progress towards a patient-specific transdiagnostic approach. Exp Neurol. 2017;287:461–72.
18. Lo M-C, Widge AS. Closed-loop neuromodulation systems: next-generation treatments for psychiatric illness. Int Rev Psychiatry. 2017;29:191–204.
19. Posternak MA, Zimmerman M. Is there a delay in the antidepressant effect? A meta-analysis. J Clin Psychiatry. 2005;66(2):148–58.
20. Brudney D, Lantos J. Agency and authenticity: which value grounds patient choice? Theor Med Bioeth. 2011;32:217–27.
21. Mackenzie R. Authenticity versus autonomy in choosing the new me: beyond IEC and NIEC in DBS. AJOB Neurosci. 2014;5:51–3.
22. Kraemer F. Authenticity or autonomy? When deep brain stimulation causes a dilemma. J Med Ethics. 2013;39:757–60.
23. Sharp D, Wasserman D. Deep brain stimulation, historicism, and moral responsibility. Neuroethics. 2016;9:173–85.
24. Berg JW. Constructing competence: formulating standards of legal competence to make medical decisions. Rutgers Law Rev: 53.
25. Bublitz JC, Merkel R. Autonomy and authenticity of enhanced personality traits. Bioethics. 2009;23:360–74.
26. Baylis F. "I am who I am": on the perceived threats to personal identity from deep brain stimulation. Neuroethics. 2011:1–14.
27. Ahlin J. What justifies judgments of inauthenticity? HEC Forum. 2018;30:361–77.
28. Erler A, Hope T. Mental disorder and the concept of authenticity. Philos Psychiatry Psychol. 2014;21:219–32.
29. Sjöstrand M, Juth N. Authenticity and psychiatric disorder: does autonomy of personal preferences matter? Med Health Care Philos. 2014;17:115–22.

30. Choi KS, Riva-Posse P, Gross RE, Mayberg HS. Mapping the "depression switch" during intraoperative testing of Subcallosal cingulate deep brain stimulation. JAMA Neurol. 2015;72:1252–60.
31. Klein E, Brown T, Sample M, Truitt AR, Goering S. Engineering the brain: ethical issues and the introduction of neural devices. Hastings Cent Rep. 2015;45:26–35.
32. Ienca M, Haselager P. Hacking the brain: brain–computer interfacing technology and the ethics of neurosecurity. Ethics Inf Technol. 2016;18:117–29.
33. Pycroft L, Boccard SG, Owen SLF, Stein JF, Fitzgerald JJ, Green AL, et al. Brainjacking: implant security issues in invasive Neuromodulation. World Neurosurg. 2016;92:454–62.
34. Denning T, Matsuoka Y, Kohno T. Neurosecurity: security and privacy for neural devices. Neurosurg Focus. 2009;27:E7.
35. Glannon W, Ineichen C. Philosophical aspects of closed-loop neuroscience. In: El Hady A, editor. Closed loop neuroscience [Internet]. San Diego: Academic Press; 2016. p. 259–70. https://books.google.com/books?hl=en&lr=&id=0ZTBCQAAQBAJ&oi=fnd&pg=PP1&dq=glannon+philosophical+aspects+of+closed-loop+neuroscience&ots=74dkrS5k9z&sig=bOJfqUkD-6FU0tJYOEdy5bQbyFw. Accessed 17 May 2017.
36. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. N Engl J Med. 2018;378:981–3.
37. Gilbert F. The burden of normality: from 'chronically ill' to 'symptom free'. New ethical challenges for deep brain stimulation postoperative treatment. J Med Ethics. 2012;38:408–12.
38. O'Brolchain F, Gordijn B. Brain–computer interfaces and user responsibility. In: Grübler G, Hildt E, editors. Brain-computer-interfaces in their ethical, social and cultural contexts [Internet]. Dordrecht: Springer; 2014. p. 163–82. Accessed 24 May 2019. https://doi.org/10.1007/978-94-017-8996-7_14.
39. Klein E. Models of the patient-machine-clinician relationship in closed-loop machine Neuromodulation. In: van Rysewyk SP, Pontier M, editors. Machine medical ethics [Internet]. Cham: Springer International; 2015. p. 273–90. Accessed 4 Nov 2019. https://doi.org/10.1007/978-3-319-08108-3_17.