Mind Control: Frontiers in Guiding the Mind

John D. Medaglia [*], **Perry Zurn** [†] [‡] **Walter Sinnott-Armstrong** [§] **Danielle S. Bassett** [¶] [‖]

[*]Department of Psychology, University of Pennsylvania, Philadelphia, PA, 19104 USA,[†]Center for Curiosity, University of Pennsylvania, Philadelphia, PA, 19104 USA,[‡]Department of Philosophy, American University, Washington, DC, 20016 USA,[§]Department of Philosophy and Kenan Institute for Ethics, Duke University, Durham, NC, 27708 USA,[¶]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, 19104 USA, and [‖]Department of Electrical & Systems Engineering, University of Pennsylvania, Philadelphia, PA, 19104 USA

The human brain is a complex network that supports mental function. The nascent field of network neuroscience applies tools from mathematics to neuroimaging data in the hopes of shedding light on cognitive function. A critical question arising from these empirical studies is how to modulate a human brain network to treat cognitive deficits or enhance mental abilities. While historically a number of tools have been employed to modulate mental states (such as cognitive behavioral therapy and brain stimulation), theoretical frameworks to guide these interventions – and to optimize them for clinical use – are fundamentally lacking. One promising and as-yet underexplored approach lies in a subdiscipline of engineering known as *network control theory*. Here, we posit that network control fundamentally relates to mind control, and that this relationship highlights important areas for future empirical research and opportunities to translate knowledge in practical domains. We clarify the conceptual intersection between neuroanatomy, cognition, and control engineering in the context of network neuroscience. Finally, we discuss the challenges, ethics, and promises of mind control.

brain network | controllability | network science | diffusion tractography | cognitive control

Introduction

Mind control is a common plot device in many genres of fiction. Its ubiquity is perhaps unsurprising: the prospect of the explicit, full control of the mind evokes alluring and startling possibilities. Fictional mind control often takes implausible forms: telepathy, magical interventions, and nefarious schemes of authoritarian organizations. In more biologically-inspired plot lines, mind control is affected by devices implanted in the subject's brain, such as in *The Matrix*: these devices manipulate neurophysiological processes resulting in a change of mental state.

In reality, mind control encompasses numerous means for influencing the mind. This includes effects mediated through the senses. Sense-mediated effects can include positive influences, such as updating one's beliefs based on presented evidence, or "nudging" someone to make healthy decisions via environmental manipulation [1]. They can also be more insiduous, such as in propaganda or brain washing. Beyond social and environmental means, mind control can also result from direct neural stimulation. Neural stimulation can include subtle modulation via pharmacological agents, or more direct manipulations with brain stimulation that result in neural discharges. The last few decades have seen a steady increase in the use of implanted devices to assist individuals with major mental disorders. The ubiquitous nature of social forms of control and increasing prevalence of neural devices motivates important questions about the control of brain processes and, by extension, mental functions.

Developing a true science of mind control could benefit from the engineering discipline known as "control theory", which addresses the question of how to guide complex systems from one state to another [2]. As a commonplace example, control systems in a modern airliner ensure that the aircraft stays aloft by automatically adjusting the plane's pitch, roll, and yaw to compensate for the turbulence in the air [3]. Like a plane, the brain is a physical system that is characterized by specific states: in this case, patterns of neural activity. The control or guidance of the brain from one state to another can either be either intrinsic (the brain controls itself [4]) or extrinsic (the brain is guided externally, for example by brain stimulation [5]).

Beyond these naturalistic forms of mind control, stimulation-based interventions have been developed for clinical cohorts, offering a powerful link between engineering and neuroscience. Indeed, the use of control theoretic approaches are particularly prevalent in the booming field of neuroprosthetics [6, 7], which can be used to treat motor disorders such as Parkinson's disease [8]. Here, control theory dictates energetically efficient strategies for deep brain stimulation to regulate motor functions mediated by subcortical areas. Similar efforts have been developed to treat obsessive compulsive disorder [9] and depression [10], suggesting their utility for deficits in the brain's cognitive control and reward circuitry. Moreover, neural control is administered through noninvasive yet powerful techniques such as transcranial magnetic stimulation [11].

To better describe how control theory from engineering can inform our understanding of both intrinsic and extrinsic forms of mind control, it is useful to invoke a network approach [4]. In a network perspective, neural components (such as individual neurons or entire brain regions) are treated as network nodes and connections between these components are treated as network edges [12, 13, 14]. By using this conceptual framework, as well as the mathematical formalism that accompanies it, we can explicitly study how control energy injected into one brain component can impact the activity in the rest of the network [4, 15]. Indeed, the network formalism allows us to capitalize on recent advances in the control of networked systems [16] which may prove useful in developing principled strategies that affect cognitive function, and the mind more generally. In particular, these strategies can theoretically guide neural systems, and thus cognition, toward target states.

In this paper, we aim to briefly review recently developed concepts and bodies of literature pertinent to a formal science of mind control. First, we introduce the general notion of control in the engineer's sense, and then we discuss recent extensions of this notion to network control. We highlight the potential relationships between brain network controllability and cognitive function, and describe what a control theoretic approach to mind control implies more broadly. We close with a discussion of ethical considerations that are relevant to society as approaches to mind control develop further in the future. While we leave mathematical details to other excellent texts, we seek throughout to provide the reader with intuitions that are sufficient to consider the state of the art, promises, and challenges that lie ahead in the control of brain networks and associated mental functions.

Control for networked systems

In the broadest sense, "control" is any form of physical influence from one entity to another; in an engineering sense, control indicates using energy to move a system from an initial state to a target state. Typically, control is exerted by a controller, C, via control energy, u, to change the state of a system, (or "plant", (P), in engineering) (Fig. 1). The strategy that the controller implements may minimize the distance between the observed state of the system and the target state, often while also maximizing technical simplicity and minimizing energy [17]. For example, the plant could be an aircraft, the controller could be the aircraft's turbine engines, and the control energy induces thrust. The aircraft has sensors that monitor its state and continuously report that information to the controller, which in turn adapts its strategy to minimize the distance between the aircraft's current trajectory and the target trajectory.

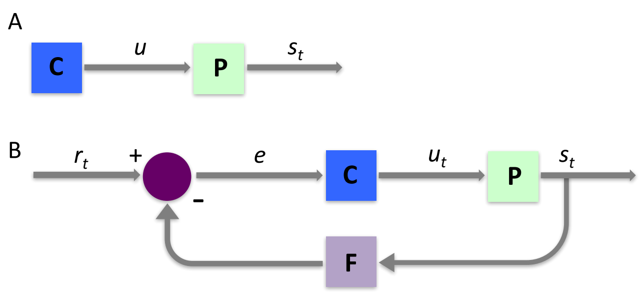

Fig. 1: **Control Theory** Generic notions of control. *(A)*: A classic open-loop control scheme. Traditionally, a controller (e.g., a system in the environment or designed device) delivers an input u to the system under control (usually termed the "plant", P) to influence the system state s_t. *(B)*: A classic closed-loop feedback control scheme. The goal is to guide the system (P) to a reference value r_t. The system state s_t is fed back through a sensor measurement F to compare to the reference value r_t. The controller C then takes the error e (difference) between the reference and the output to change the control inputs u to the system under control (P).

In the context of human cognition, the plant is the neural tissue that supports cognition – e.g., a single neuron, an ensemble of neurons, a system of brain regions, or the whole brain [18, 8, 4, 19]. For example, in deep brain stimulation in Parkinson's disease, the controller is the stimulation device, the energy is an applied current, and the system is a portion of the basal ganglia (often the globus pallidus or putamen). The neural reference state is often represented in the frequency domain, and the neural control goal is to achieve a target state of basal ganglia activity – and the rest of the motor system with which it interacts – that facilitates unimpaired motor movements [8, 20]. Beyond simple motor function, neural control can be used to treat obsessive compulsive disorder, where patients engage in repetitive behaviors following overwhelming urges to do so. Deep brain stimulation to subcortical circuitry effectively reduces symptoms of the disorder and substantially improves quality of life [21], putatively by normalizing neural activity in a fronto-striatal circuit [9].

One of the greatest challenges in applying ideas from control theory to the human brain is that input or stimulation to one area does not affect only that area. Instead, input to a single area sets off a cascade of changes that distribute influences throughout the rest of the brain in a manner that is difficult to characterize [5]. Progress in understanding this complexity has been hampered by the lack of a mathematical approach and finely resolved network data to deal with the pattern of interconnections between brain areas through which energy can propagate [22]. In light of this challenge, the recent development of the field of *connectomics* is particularly promising [23]. Here brain regions (or individual or groups of neurons) are treated as network nodes, and the interconnections between brain regions or neurons are treated as network edges. Such a network representation can directly inform our application of control theory to the brain [4]. Indeed, as we will discuss in greater detail throughout this paper, it may offer a much needed framework for the use of noninvasive brain stimulation techniques such as transcranial magnetic, direct current, and alternating current stimulation to influence cognitive functions in health [24] and disease [25, 26, 27].

The study of how to control a complex network is commonly referred to as *network control theory* [28, 29]. A networked system is represented as a "graph" $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of interconnected elements, where $\mathcal{V}$ is a vertex (or node) set and $\mathcal{E}$ is an edge set. We store this information in an adjacency matrix A, whose ij^{th} element indicates the strength of the edge between node i and node j. Nodes represent components of the system (e.g., neurons or brain regions), and edges represent connections or interactions between nodes (e.g., axons or white matter tracts). Network control theory is the study of how to design control inputs to a network that can be used to guide the system from an initial state to a target state (see Fig. 2) [15, 30]. Because energy propagates through discrete structural connections, network control is conceptually appropriate for studying how to affect the mind by manipulating the energy in specific parts of the brain, thereby inducing dynamic trajectories: changes in neural states over time that support cognitive functions. This highlights an important dual nature to mind control: the *neural* control goal is to influence *neural* states from one to another. The *psychological* control goal presumably depends on neural states, but can be sensibly discussed in its own terms. Indeed, if no psychological, physical, or social consequences were associated with depression, there would be no such thing as a mind control goal for depression.

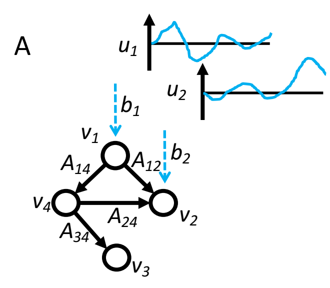

Fig. 2: **Network control.** *(A)*: Network nodes v are connected through edges A_{ij}. Control energy u can be administered to nodes via inputs b. *(B)* Control input can be designed to guide a network from an initial state (s_0) to a target state later in time (s_t). Figure adapted with permission from [3].

Brain controllability and guiding the mind

In network control theory applied to the brain, mathematical tools can be applied in any case including structural pathways in the brain (e.g., individual axons or white matter bundles) that connect neural elements (e.g., single neurons or groups of neurons) to one another. Initial applications of network control theory to the human connectome suggest that the brain may employ distinct control strategies to guide mental processes [4]. Three well-known cognitive systems each display different patterns of structural connectivity as estimated by white matter tractography, and those patterns facilitate different types of control. For example, the fronto-parietal system is a set of regions known to be strongly coherent at rest and to facilitate the human's ability to switch between different tasks [31, 32]. Interestingly, this system is relatively sparsely structurally connected with the rest of the brain. Results from theoretical network control theory applied to this system suggest that the fronto-parietal system is optimized for moving the brain into difficult-to-reach states [4] along an "energy landscape", which defines the possible states and transitions of the network [33]. Dorsal and ventral attention systems [34, 35] are known to be neither sparsely nor densely connected with the rest of the brain, and are instead predicted to be optimized for integrating or segregating other parts of the brain. Finally, the brain's baseline system – the so-called "default mode" [36, 37, 38] – is strongly structurally connected with the brain, and these patterns of structural connections are thought to facilitate the default mode's role in driving the brain to many easy-to-reach states. These findings suggest that the brain is organized into structurally distinct control systems [4].

Based on this previous work, it is intuitively plausible that distinct brain regions could be identified as candidate targets for interventions that can influence brain dynamics into distinct dynamic trajectories. Supporting such a notion, evidence from numerical simulations suggests that the structural connections emanating from a brain region directly impact the transmission of stimulation from the targeted area to the rest of the brain [5]. In that study, stimulation to regions in the default mode imparted large global change in brain activity, suggesting the importance of considering individual differences in white matter tracts in brain stimulation protocols. In addition to being a potential target for brain stimulation, regions in the default mode may also play a role in homeostasis following stimulation, as evidence suggests that they are the least energetically costly target state [15].

While these initial studies have focused on large-scale connectomes in the human, it is important to note that network control theory can be applied to other species and other spatial scales. As we will describe in greater detail in the next section, network control theory might be able to offer insights into the computations that occur in single neurons [39]. For the level of organization associated with many cognitive functions, neural control goals can focus on affecting single neurons to – likely subtly – influence cognition. In addition, control goals could involve neuronal ensembles [40, 41, 42], and large-scale distributed neural circuits [43, 44, 45]. Presumably, cognition – and arguably its control – occurs over multiple spatial resolutions [46], offering distinct targets for network control in support of mental function (see Fig. 3). The use of neuroimaging continues to be informative regarding the macro scale organization of networks that support distinct cognitive processes [47]. However, to gain increasing control over cognitive function, a more fundamental characterization of cognition will be required. Specifically, in emerging theoretical work on *neural* control across the structural connectome, the mapping between brain dynamics and specific cognitive processes will be critical to inform *psychological* (i.e., "mind") control.

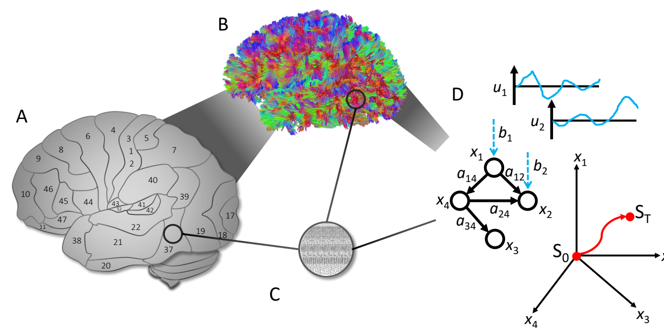

Fig. 3: **Brain control.** *(A)*: The gross anatomical organization of the brain can be described by cytoarchitectonic regimes that are thought to serve distinct roles in neural computation [48]. *(B)*: Modern imaging techniques such as diffusion weighted imaging can provide information about the macro-scale connectivity among brain regions (the "connectome" [23]). *(C)*: Low level cellular organization facilitates information processing and is embedded within the macro-scale connectome. *(D)*: The structural and dynamic organization of the brain at multiple scales can be represented as a networked system that can be guided using control energy targeted to specific brain regions.

Controlling specific mental functions

In reality, exogenous control of mental functions would require not just theoretical tools from network control theory, but a marriage of these tools with our current knowledge of the neural processes enabling cognition: namely, neural codes. Neural codes occur in several forms. These are typically thought to involve *temporal* characteristics of neural firing within a certain *population* of neurons [49]. The population may include relatively few neurons to increase information processing capacity of a neural system [50]. Temporal rate codes can be represented in the frequency domain, where information is represented in how quickly neurons fire [51, 52]. Temporal codes based on spike timing encode information in the delay between a stimulus and neural discharge [53]. Temporal codes can be multiplexed to increase encoding capacity, disambiguate stimuli, and stabilize representations [39] (see Fig. 4 for a schematic example in visual perception).

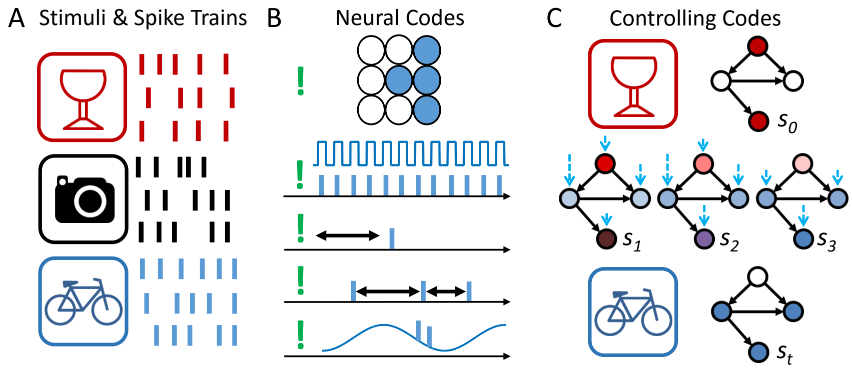

Fig. 4: **Neural codes and cognition.** A speculative schematic for the potential control of object perception. *(A)*: Neurons transmit information in the form of neural discharges or "spikes" associated with stimuli. Here, neural spike trains in distinct colors represent different stimuli. *(B)*: Neural spike trains can be analyzed to determine the nature of coding that supports distinct representations and processes that constitute stimuli. The green exclamation point represents an arbitrary stimulus that can (in principle) be represented by a number of possible codes. Top to bottom: a stimulus representation is maintained by a population of neurons, which may use frequency or "rate" coding, latency coding, interspike interval codes, or the phase of firing. *(C)*: Top: the wine glass is represented by the initial neural state s_0 of two neurons in the four neuron system. Middle to bottom: control energy is applied to different neurons in varying quantities over time to induce a transition to a state representing the bicycle. s_t represents the state of the neural network at time t. Realizing the control strategy requires the right neural population, code, and manipulating apparatus. While we select a visual perception example for clarity, a similar intuition can apply to any neurocognitive process that involves temporal codes in populations of neurons.

How neural codes relate to cognitive states and transitions between states is not yet fully understood. While the nature and relevance of cognitive "representations" remains an open issue [54], a well-formed control strategy includes a well defined neural control goal and a mapping from neural state to cognitive representations and processes. To move cognitive neuroscience from correlation to causative models, a combination of direct manipulation and observation are required. While speculative at this point in scientific history, it is possible that one could combine an understanding of neural codes, experimental manipulation via brain stimulation, and network control theory to design control strategies to achieve certain mental states. This defines the frontier of mind control.

Technologies for control

Some notable practical contexts for mind control exist in nascent stages of development that could benefit from analysis within network control theory. In clinical contexts, network control theory could inform strategies to re-establish preclinical function. In addition, it could improve prior function in the case of cognitive enhancement. At present, a number of technologies for neural stimulation exist and are routinely used in experimental and clinical neuroscience (Fig. 4). At present, techniques typically capitalize on different spatiotemporal properties of electromagnetism. Microstimulation can influence the activity of single neurons. Most techniques used in cognitive neuroscience and clinics operate at a scale much coarser than a single neuron. For example, transcranial magnetic stimulation non-invasively induces current in cortical tissue at a right angle to a magnetic field that passes through

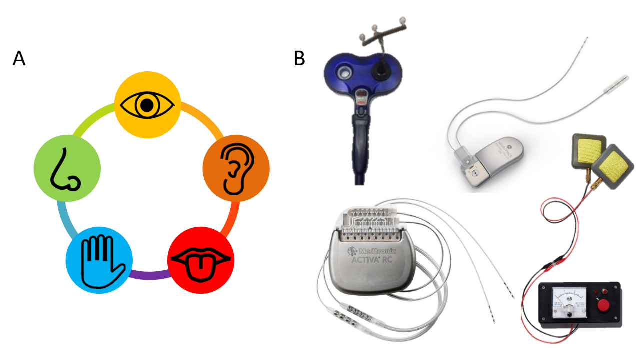

Fig. 5: **Forms of control.** *(A)*: While not the primary emphasis of the current discussion, we note that one form of neural control is constantly mediated through sensation, which influences immediate experience and learning and the basis of psychological treatments such as cognitive behavioral therapy *(B)*: Neural control can be administered via direct noninvasive or invasive neural stimulation devices. Clockwise from top left: MagStim$^{\text{TM}}$ transcranial magnetic stimulation coil, NeuroPace$^{\text{TM}}$ implanted stimulator, custom transcranial direct current stimulator, and MedTronic$^{\text{TM}}$ deep brain stimulator. Images copyrighted by devices manufacturers.

the cranium into the brain's gray matter, where it causes action potentials in large populations of neurons. With much less spatial precision, approaches such as transcranial direct current stimulation use current administered to the scalp with the goal of modulating neural firing thresholds in the cortex with. Invasive approaches such as deep brain stimulation applies implanted electrodes to influence local field potentials in groups of neurons. Evidence suggests that dynamics supporting cognitive function can be influenced by these techniques at a low energetic and risk cost [55].

Most applications of noninvasive electrical and magnetic stimulation are a form of open-loop control without feedback about internal neural states. This stands in contrast to closed-loop control, which involves feedback about internal neural states [8]. In open-loop control, control strategies are often evaluated based on their success in influencing specific cognitive and behavioral functions [56] or their efficacy in influencing clinical outcomes [57]. Here, the goal or "reference" state is the cognitive status of the subject rather than the internal neural states. Studies in this area have elucidated the putative localizations for specific functions in the brain and led to FDA-approved therapies for depression, which is characterized by significant emotional and cognitive disturbances [58]. This type of research continues to provide substantial scientific and clinical benefits in behavioral neuroscience in the absence of true closed-loop control with neural measurement feedback. It is also an example for how a *psychological* control goal can be approached even when knowledge of specific neural codes is limited or completely absent.

In closed-loop control, a critical component in the controller-observer (feedback sensor) cycle is the ability to detect states of the system. Devices such as deep brain stimulation benefit from real-time detection of neural states, where the control strategy can update and adapt its input sequence based on the current states of the system [8]. At these scales, implanted microsensors record the local activity in neurons [59]. At higher spatiotemporal scales, spatially coarse states can be measured by techniques such as EEG [60], and temporally coarse states can be measures with real-time fMRI [61]. In biofeedback paradigms, the participant can observe his or her own measured neural states, and can adaptively learn to control these signals [62].

Good enough control

To truly understand the means to control the mind suggests that one has sufficiently observed it [63]. This is no simple task. Brain stimulation strategies for mind control are limited by knowledge of the underlying brain network architecture, the nature of neural codes, and the limitations of technology and computing [8]. Given the difficulties inherent in measuring neural states, modeling system dynamics, and estimating signals within noise, simplifying assumptions can be useful in facilitating insights into intrinsic [4] and extrinsic [5] control of brain function. Using these simplifying assumptions, studies suggest that in fact the brain is very difficult, if not impossible, to control [4, 64]. Thus, in truth, one can likely only "guide" the brain in part rather than "control" it in its entirety.

Encouragingly, neural control strategies can be designed in the context of limited information about a neural system [8]. The efficacy of limited control strategies can be evaluated in guiding the mind where information about the representation of cognition is available. Various techniques in neuroscience provide information about neural states with varying degrees of spatiotemporal precision. To date, efforts in mapping cognitive functions within the brain have relied on techniques ranging from invasive microstimulation and recording on limited sets of neurons to non-invasive imaging and computational techniques such as multivoxel pattern analysis [65]. In addition, transitions between states at a coarse scale can be studied using dynamic network approaches [44, 66, 45]. Thus, while an ideal control strategy would include precise mapping between neural codes and the mind, numerous findings in brain stimulation research at multiple scales of neural organization suggest that the mind can be influenced with interpretive and practical value in psychological control. Findings from neuroimaging suggest that cognitively relevant information can be identified

In particular, guiding the mind is perhaps most pragmatic in clinical scenarios where internal neural states are not directly measured, sensor feedback is difficult to maintain, and costs for comprehensive evaluation are prohibitive. For these types of scenarios, engineers are developing tools to influence the dynamic trajectory of the system using control strategies with limited access to the system and finite control energy [67]. More specifically, these techniques offer guidelines to steer the network into the intended reference state via cost-efficient strategies that both directly influence the system and allow the system's natural dynamics to help drive the system to the control goal. Future studies utilize these and similar approaches to identify candidate strategies to guide brain network dynamics in the context of missing information and inadequate control.

To close this speculative discussion, we mention a few additional limitations. First, we motivate these ideas based on the notion of "structural controllability": the control input is mediated through the nodes (brain regions) in the network, and its influences are relayed through edges (white matter tracts) to other nodes. In principle, other control strategies can be designed [68], including (i) modifying network edges by introducing bypasses for damaged tissue via neuroprosthesis [69] and (ii) modifying local regional dynamics using pharmacology [70]. Second, we note that mind control holds no special ontological status relative to motor control. Indeed, the distinction between motor and cognitive processes may simply be an artifact of historical emphasis [71]. If both motor and cognitive function are fundamentally computational processes [72], both can be informed by neural control engineering. However, there may be differences in the quality and degree of control required across these domains, as well as different stakes.

The ethics of brain control

As efforts to guide complex brain processes advance, we will not only need new theoretical and technical tools. We will also face new societal, legal, and ethical challenges. Our best chance of meeting those challenges is through ongoing, rigorous discussion between scientists, ethicists, and policy makers.

Even at this early stage in the development of the science of mind control, present and future ethical issues are important to consider. These issues pertain to mind control research as well as to clinical applications of mind control as they change and develop. In both cases, it is crucial to maximize benefits to society and protect against harm. In addition, ethical restrictions that apply in the use of mind control to treat dysfunction and alleviate suffering may differ from those that apply where mind control may be used to enhance typical function [73].

Insofar as the practices of mind control have and will be undertaken in experimental and clinical contexts, the four basic principles of medical and research ethics apply here: nonmaleficence, beneficence, justice, and autonomy [74]. Adhering to these principles is a first step toward ensuring that efforts to guide the mind enhance human welfare without violating human rights. Here we will take direct neural and magnetic stimulation as our primary points of analysis. In addition, we will identify basic ethical issues that may be important to the more complex case of sensation-mediated stimulation, which includes the broader social and environmental forms of control that individuals experience in daily living.

Non-maleficence. The principle of nonmaleficence should supersede any implementation of control in experimental or clinical contexts. In the development and application of these techniques, both the physical and psychological safety of subjects should be considered. For neural control devices, deep brain stimulation involves implanted neural electrodes and is associated with a risk of incidental neural damage and infection [75]. Psychiatric side effects include depression, delusion, euphoria, and disinhibition [76]. Transcranial magnetic stimulation, on the other hand, is noninvasive and carries an exceptionally rare risk of seizure [77]. As these and other technologies are refined and created, improved safeguards against tissue damage and adverse psychological effects must remain a high priority. In current practice, none of these harms is intended either as a goal or as a means; researchers are therefore not directly maleficent, but may nevertheless be indirectly responsible for their subjects suffering. Of course, it is possible that mind control will be used for nefarious ends either by individual evil-doers or as part of an oppressive system. As the potential for mind control is explored, therefore, efforts should be made to reduce the potential for indirect maleficence and safeguard against the possibility of direct maleficence.

Beneficence. In addition to preventing harm, society should maintain a premium on increasing individual and group welfare via mind control. In applied brain stimulation, clinical symptoms can be reduced and improvements in cognitive performance in healthy individuals can be produced [73]. This may include new and more powerful approaches to clinical symptom reduction and more robust cognitive enhancement [73]. As theoretical models of brain function at multiple scales develop, new opportunities for maximally beneficent control should be conceptualized and tested. Nonetheless, given early misuses of lobotomies and electroconvulsive therapy, it is important to separate uses of mind control to improve mental function in clinical cases from its potential uses for social normalization or moral correction, such as 'treatments' of selfishness, homosexuality, or criminality.

Justice. The principle of justice demands that the application of mind control neither participate in nor exacerbate systems of inequality or exploitation. As such, already existing financial barriers to neural control based treatments and enhancements should be reduced. In turn, access to information about the benefits and risks of such applications should be judiciously increased, in concert with efforts to improve general public health education, especially in underserved communities and countries. To date, mind control application exists in relatively circumscribed clinical and experimental contexts. If, however, mind control strategies for optimal cognitive function were to become more widely available, or even marketable for public consumption, safeguards should be put in place against the construction of an *enhanced* class, which may result in harm to the *unenhanced* in competitive environments [73].

Autonomy. In the process of guiding the mind, it is imperative that the individual's autonomy, or power of self-determination, remains intact. Autonomy is violated by any involuntary use of mind control on anyone, whether by explicit or implicit coercion. Explicit coercion refers to forcing individuals to undergo applied mind control against their conscious will, whether for the perceived greater good of society or the advancement of some constituency. The strongest safeguard against explicit coercion to date has been informed consent [78]; however, the informed consent process still needs to be improved [79] and additional methods to supplement it should be developed. This is especially relevant to potential applications of mind control in vulnerable populations. In the future, moreover, should a group of enhanced individuals gain dominance, unenhanced individuals could conceivably be at risk of involuntary medicalization [73].

Implicit coercion refers to manipulating individuals through the activation of external or internal pressure rather than force. While implicit coercion has sullied the history of research ethics [74], the distinction between implicit coercion and participatory incentive today still needs further clarification. For mind control, perhaps the surest source of implicit coercion will be the social pressures to increase productivity or efficiency in competitive environments [80, 81, 82, 83]. Importantly,there is no definitive answer to the limits to autonomy. The reality of structural constraints and mediated freedoms makes appropriate levels of adult autonomy difficulty to calibrate. Nevertheless, research clinicians should make an effort not to capitalize on social conditions that might introduce implicit coercion. For example, poverty may predispose subjects to enter paid research trials, leading to higher rates of experimentation on lower class bodies. As such, the potential for subjects to enroll in research and clinical designs under conditions of implicit coercion should be the focus of redoubled ethical review.

While our discussion here bears specifically on humans with a typical adult capacity for self-determination, further ethical considerations would come into play were mind control to be available for people lacking such capacity – for example, children, some elderly individuals, and adults with certain forms of mental or physical disabilities.

Rethinking human persons

As mind control develops as an area of research and practice, the ability to interact intelligently with human nature may bring certain stakes into sharper focus. Humans privilege the notion of a "mind" in their identity and perceived internal locus of control over the notion of simple stimulus-behavior pairings [84]. Further, within minds, humans privilege some traits such as social comfort, honesty, kindness, empathy, and fairness as more fundamental than functions such as concentration, wakefulness, and memory [85]. Indeed, these different values depend on the notion of conscious identity and are often at the core of common ethical distinctions applied to humans *versus* other animals [86]. Importantly, modern notions of human persons, influenced by continuing advances in the cognitive and brain sciences, erode the classical boundary between the ethical treatment of humans and animals [87]. These theories suggest that tolerance for mind control may scale with ethical issues identified across all animals. In kind, this implies that there exists a need to evaluate the practice of mind control on both human and non-human animals and its permissible scope in principle.

As the science of mind control advances, it will be important to clarify acceptable control practices with respect to our fundamental nature and self-identity. In addition, the potential for mind control to undermine responsibility connects to our fundamental intuitions about whether we really control what we do [88]. Implicit within this new technology, then, is the call to redefine ourselves. For this reason, scientists, clinicians, ethicists, and philosophers will need to work together.

Conclusions

The study of mind control in human brains has been developing over several years without any particular name. Here, we have described mind control to be fundamentally a problem of network control in the human brain. New frontiers at the intersection of network neuroscience and cognitive science can provide striking new questions and possibilities in understanding and controlling cognitive function. Our success will be predicated on the development on robust theories of neural coding and advances in technologies for recording and manipulation neural activity. In reclaiming a term long relegated to science fiction, new opportunities within and between computational, cognitive, clinical, ethical, and control neuroscience may produce a new era in the science of guiding the mind.

ACKNOWLEDGMENTS. JDM acknowledges support from the Office of the Director at the National Institutes of Health through grant number 1-DP5-OD-021352-01. DSB acknowledges support from the John D. and Catherine T. MacArthur Foundation, the Alfred P. Sloan Foundation, the Army Research Laboratory and the Army Research Office through contract numbers W911NF-10-2-0022 and W911NF-14-1-0679, the National Institute of Health (2-R01-DC-009209-11, 1R01HD086888-01, R01-MH107235, R01-MH107703, R01MH109520, 1R01NS099348 and R21-M MH-106799), the Office of Naval Research, and the National Science Foundation (BCS-1441502, CAREER PHY-1554488, BCS-1631550, and CNS-1626008).The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.

1. Thaler, R. H., Sunstein, C. R. & Balz, J. P. Choice architecture. The behavioral foundations of public policy (2014).
2. Kailath, T. Linear Systems (Prentice-Hall, 1980).
3. Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. Controllability of complex networks. Nature 473, 167–173 (2011).
4. Gu, S. et al. Controllability of structural brain networks. Nature communications 6 (2015).
5. Muldoon, S. F. et al. Stimulation-based control of dynamic brain networks. PLoS Computational Biology 12, e1005076 (2016).
6. Humayun, M. et al. Recent results from second sights argus® ii retinal prosthesis study. Investigative Ophthalmology & Visual Science 54, 349–349 (2013).
7. Arts, R. A., George, E. L., Stokroos, R. J. & Vermeire, K. Review: cochlear implants as a treatment of tinnitus in single-sided deafness. Current opinion in otolaryngology & head and neck surgery 20, 398–403 (2012).
8. Schiff, S. J. Neural Control Engineering (The MIT Press, 2012).
9. Figee, M. et al. Deep brain stimulation restores frontostriatal network activity in obsessive-compulsive disorder. Nature neuroscience 16, 386–387 (2013).
10. Mayberg, H. S. et al. Deep brain stimulation for treatment-resistant depression. Neuron 45, 651–660 (2005).
11. Hallett, M. Transcranial magnetic stimulation and the human brain. Nature 406, 147–150 (2000).
12. Bassett, D. S. & Bullmore, E. Small-world brain networks. The neuroscientist 12, 512–523 (2006).
13. Bullmore, E. & Sporns, O. Complex brain networks: Graph theoretical analysis of structural and functional systems. Nature Review Neuroscience 10, 186–198 (2009).
14. Bullmore, E. T. & Bassett, D. S. Brain graphs: graphical models of the human brain connectome. Annual Reviews of Clinical Psychology 7, 113–140 (2011).
15. Betzel, R. F., Gu, S., Medaglia, J. D., Pasqualetti, F. & Bassett, D. S. Optimally controlling the human connectome: the role of network topology. Scientific Reports In Press (2016).
16. Pasqualetti, F., Zampieri, S. & Bullo, F. Controllability metrics, limitations and algorithms for complex networks. IEEE Transactions on Control of Network Systems 1, 40–52 (2014).
17. Lee, E. B. & Markus, L. Foundations of optimal control theory. Tech. Rep., DTIC Document (1967).
18. Stigen, T., Danzl, P., Moehlis, J. & Netoff, T. Controlling spike timing and synchrony in oscillatory neurons. BMC Neuroscience 12, P223 (2011).
19. Nabi, A. & Moehlis, J. Single input optimal control for globally coupled neuron networks. Journal of neural engineering 8, 065008 (2011).
20. Sarma, S. V. et al. Using point process models to compare neural spiking activity in the subthalamic nucleus of parkinson's patients and a healthy primate. Biomedical Engineering, IEEE Transactions on 57, 1297–1305 (2010).
21. Greenberg, B. D. et al. Three-year outcomes in deep brain stimulation for highly resistant obsessive–compulsive disorder. Neuropsychopharmacology 31, 2384–2393 (2006).
22. Gu, S. et al. The energy landscape of neurophysiological activity implicit in brain network structure. arXiv 1607, 01959 (2016).
23. Sporns, O., Tononi, G. & Kötter, R. The human connectome: a structural description of the human brain. PLoS Comput Biol 1, e42 (2005).
24. Miniussi, C., Harris, J. A. & Ruzzoli, M. Modelling non-invasive brain stimulation in cognitive neuroscience. Neuroscience & Biobehavioral Reviews 37, 1702–1712 (2013).
25. Brunoni, A. R. et al. Clinical research with transcranial direct current stimulation (tdcs): challenges and future directions. Brain stimulation 5, 175–195 (2012).
26. Teneback, C. C. et al. Changes in prefrontal cortex and paralimbic activity in depression following two weeks of daily left prefrontal tms. The Journal of neuropsychiatry and clinical neurosciences (2015).
27. Nitsche, M. tdcs/tacs and plasticity. Clinical Neurophysiology 127, e32 (2016).
28. Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. Controllability of complex networks. Nature 473, 167–173 (2011).
29. Ruths, J. & Ruths, D. Control profiles of complex networks. Science 343, 1373–1376 (2014).
30. Gu, S. et al. Optimal trajectories of brain state transitions. arXiv 1607, 01706 (2016).
31. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. Annual review of neuroscience 24, 167–202 (2001).
32. Dosenbach, N. U. et al. Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci U S A 104, 11073–11078 (2007).
33. Watanabe, T., Masuda, N., Megumi, F., Kanai, R. & Rees, G. Energy landscape and dynamics of brain activity during human bistable perception. Nature communications 5 (2014).
34. Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3, 201–215 (2002).
35. Sestieri, C., Corbetta, M., Spadone, S., Romani, G. L. & Shulman, G. L. Domain-general signals in the cingulo-opercular network for visuospatial attention and episodic memory. J Cogn Neurosci 26, 551–568 (2014).
36. Raichle, M. E. et al. A default mode of brain function. Proceedings of the National Academy of Sciences 98, 676–682 (2001).
37. Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L. & Raichle, M. E. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. Proceedings of the National Academy of Sciences 103, 10046–10051 (2006).
38. Fox, M. D. & Raichle, M. E. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nature Reviews Neuroscience 8, 700–711 (2007).
39. Panzeri, S., Brunel, N., Logothetis, N. K. & Kayser, C. Sensory neural codes using multiplexed temporal scales. Trends in neurosciences 33, 111–120 (2010).
40. Sutskever, I., Vinyals, O. & Le, Q. V. Sequence to sequence learning with neural networks. In Advances in neural information processing systems, 3104–3112 (2014).
41. Merchant, H., Harrington, D. L. & Meck, W. H. Neural basis of the perception and estimation of time. Annual review of neuroscience 36, 313–336 (2013).
42. Stolier, R. M. & Freeman, J. B. Neural pattern similarity reveals the inherent intersection of social categories. Nature neuroscience (2016).
43. Bassett, D. S., Yang, M., Wymbs, N. F. & Grafton, S. T. Learning-induced autonomy of sensorimotor systems. Nat Neurosci 18, 744–751 (2015).
44. Calhoun, V. D., Miller, R., Pearlson, G. & Adalı, T. The chronnectome: time-varying connectivity networks as the next frontier in fmri data discovery. Neuron 84, 262–274 (2014).
45. Medaglia, J. D., Lynall, M. E. & Bassett, D. S. Cognitive network neuroscience. J Cogn Neurosci 27, 1471–1491 (2015).
46. Bassett, D. S. & Gazzaniga, M. S. Understanding complexity in the human brain. Trends in cognitive sciences 15, 200–209 (2011).
47. Medaglia, J. D., Lynall, M.-E. & Bassett, D. S. Cognitive network neuroscience. Journal of cognitive neuroscience (2015).
48. Brodmann, K. Vegleichende Lokalisationslehre der Grosshirnde (Barth, 1909).
49. Shamir, M. Emerging principles of population coding: in search for the neural code. Current opinion in neurobiology 25, 140–148 (2014).
50. Zaslaver, A. et al. Hierarchical sparse coding in the sensory system of caenorhabditis elegans. Proceedings of the National Academy of Sciences 112, 1185–1189 (2015).
51. Lisman, J. E. & Jensen, O. The theta-gamma neural code. Neuron 77, 1002–1016 (2013).
52. Miura, K., Mainen, Z. F. & Uchida, N. Odor representations in olfactory cortex: distributed rate coding and decorrelated population activity. Neuron 74, 1087–1098 (2012).
53. Ainsworth, M. et al. Rates and rhythms: a synergistic view of frequency and temporal coding in neuronal networks. Neuron 75, 572–583 (2012).
54. Van Gelder, T. What might cognition be, if not computation? The Journal of Philosophy 92, 345–381 (1995).
55. Barker, A. T., Jalinous, R. & Freeston, I. L. Non-invasive magnetic stimulation of human motor cortex. The Lancet 325, 1106–1107 (1985).
56. Silvanto, J., Muggleton, N. & Walsh, V. State-dependency in brain stimulation studies of perception and cognition. Trends in cognitive sciences 12, 447–454 (2008).

57. Fox, M. D. et al. Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and neurological diseases. Proceedings of the National Academy of Sciences 111, E4367–E4375 (2014).
58. George, M. S. & Aston-Jones, G. Noninvasive techniques for probing neurocircuitry and treating illness: vagus nerve stimulation (vns), transcranial magnetic stimulation (tms) and transcranial direct current stimulation (tdcs). Neuropsychopharmacology 35, 301–316 (2010).
59. Kozai, T. D. Y. et al. Ultrasmall implantable composite microelectrodes with bioactive surfaces for chronic neural interfaces. Nature materials 11, 1065–1073 (2012).
60. Mullen, T. R. et al. Real-time neuroimaging and cognitive monitoring using wearable dry eeg. Biomedical Engineering, IEEE Transactions on 62, 2553–2567 (2015).
61. Zotev, V., Phillips, R., Yuan, H., Misaki, M. & Bodurka, J. Self-regulation of human brain activity using simultaneous real-time fmri and eeg neurofeedback. NeuroImage 85, 985–995 (2014).
62. Florin, E., Bock, E. & Baillet, S. Targeted reinforcement of neural oscillatory activity with real-time neuroimaging feedback. Neuroimage 88, 54–60 (2014).
63. Kalman, R. E. A new approach to linear filtering and prediction problems. Journal of basic Engineering 82, 35–45 (1960).
64. Tarokh, M. Measures for controllability, observability and fixed modes. Automatic Control, IEEE Transactions on 37, 1268–1273 (1992).
65. Norman, K. A., Polyn, S. M., Detre, G. J. & Haxby, J. V. Beyond mind-reading: multi-voxel pattern analysis of fmri data. Trends in cognitive sciences 10, 424–430 (2006).
66. Kopell, N. J., Gritton, H. J., Whittington, M. A. & Kramer, M. A. Beyond the connectome: The dynome. Neuron 83, 1319–1320 (2014).
67. Motter, A. E. Networkcontrology. Chaos: An Interdisciplinary Journal of Nonlinear Science 25, 097621 (2015).
68. Nowzari, C., Preciado, V. M. & Pappas, G. J. Analysis and control of epidemics: A survey of spreading processes on complex networks. IEEE Control Systems 36, 26–46 (2016).
69. Bouton, C. E. et al. Restoring cortical control of functional movement in a human with quadriplegia. Nature 533, 247–250 (2016).
70. Zhao, S. & Iyengar, R. Systems pharmacology: network analysis to identify multiscale mechanisms of drug action. Annual review of pharmacology and toxicology 52, 505 (2012).
71. Rosenbaum, D. A. The cinderella of psychology: the neglect of motor control in the science of mental life and behavior. American Psychologist 60, 308 (2005).
72. Churchland, P. S. & Sejnowski, T. J. Neural representation and neural computation. (The MIT Press, 1989).
73. Hamilton, R., Messing, S. & Chatterjee, A. Rethinking the thinking cap ethics of neural enhancement using noninvasive brain stimulation. Neurology 76, 187–193 (2011).
74. Shamoo, A. E. & Resnik, D. B. Responsible conduct of research (Oxford University Press, 2009).
75. Bronstein, J. M. et al. Deep brain stimulation for parkinson disease: an expert consensus and review of key issues. Archives of neurology 68, 165–165 (2011).
76. Pinsker, M., Amtage, F., Berger, M., Nikkhah, G. & van Elst, L. T. Psychiatric side-effects of bilateral deep brain stimulation for movement disorders. In Stereotactic and Functional Neurosurgery, 47–51 (Springer, 2013).
77. Been, G., Ngo, T. T., Miller, S. M. & Fitzgerald, P. B. The use of tdcs and cvs as methods of non-invasive brain stimulation. Brain research reviews 56, 346–361 (2007).
78. Wear, S. & Moreno, J. D. Informed consent: patient autonomy and physician beneficence within clinical medicine. In HEC Forum, vol. 6, 323–325 (Springer, 1994).
79. Lentz, J., Kennett, M., Perlmutter, J. & Forrest, A. Paving the way to a more effective informed consent process: Recommendations from the clinical trials transformation initiative. Contemporary Clinical Trials 49, 65–69 (2016).
80. McCabe, S. E., Knight, J. R., Teter, C. J. & Wechsler, H. Non-medical use of prescription stimulants among us college students: Prevalence and correlates from a national survey. Addiction 100, 96–106 (2005).
81. Sahakian, B. & Morein-Zamir, S. Professor's little helper. Nature 450, 1157–1159 (2007).
82. Fox, D. Brain buzz. Nature 472, 156–159 (2011).
83. Snowball, A. et al. Long-term enhancement of brain function and cognition using cognitive training and brain stimulation. Current Biology 23, 987–992 (2013).
84. Wilson, R. A. & Lenart, B. A. Extended mind and identity. In Handbook of neuroethics, 423–439 (Springer, 2015).
85. Riis, J., Simmons, J. P. & Goodwin, G. P. Preferences for enhancement pharmaceuticals: The reluctance to enhance fundamental traits. Journal of Consumer Research 35, 495–508 (2008).
86. Olson, E. T. The human animal: Personal identity without psychology (Oxford University Press, 1999).
87. Singer, P. Practical ethics (Cambridge university press, 2011).
88. Dennett, D. C. Elbow room: The varieties of free will worth wanting (MIT Press, 2015).