

Journal of Cognition and Neuroethics

Brain Rays, Advertising, and Fancy Suits: The Ethics of Mind Control

Brent Kiou

University of Utah

Biography

Dr. Brent Kiou is currently a psychiatry resident at the University of Utah. He completed medical school at the David Geffen School of Medicine at UCLA, and did his doctoral work in philosophy at UCLA. His research interests include applied ethics, metaethics, contract theory, and philosophy of psychiatry, and he has published on the ethics of enhancement, genetic discrimination, and the relationship between decision-making autonomy and values.

Publication Details

Journal of Cognition and Neuroethics (ISSN: 2166-5087). March, 2014. Volume 2, Issue 1.

Citation

Kiou, Brent. 2014. "Brain Rays, Advertising, and Fancy Suits: The Ethics of Mind Control." *Journal of Cognition and Neuroethics* 2 (1): 191–210.

Brain Rays, Advertising, and Fancy Suits: The Ethics of Mind Control

Brent Kiouss

Abstract

Many science-fictional kinds of mind control—technologies that might be used to manipulate other persons' thoughts and behavior—seem patently impermissible. The most natural account of this impermissibility is Kantian: mind control technologies would undermine others' rational capacities, but our duty to respect each others' rational personhood precludes this. I argue, however, that the Kantian view is inconsistent with the fact, demonstrated by several examples, that we permissibly manipulate others' thoughts and behavior in a variety of mundane ways on a regular basis. After considering possible defenses of the Kantian view, I delineate an alternative theory, according to which we can distinguish permissible from impermissible sorts of mind control on largely contractualist grounds.

Keywords

Kant, rationality, mind control, contractualism

Introduction

Mind control sounds scary. The thought that a person might use brain rays, psychosurgery, or some novel technological process to make others think and do things horrifies us. Here I will assume that this horror reflects a moral truth: the use of certain sorts of mind control—like a brain ray or a neural parasite or any of the other products of science-fiction writers' fevered imaginations—would be terribly wrong. In what follows, I will try to make sense of this fact. Though the project is more than a little whimsical, it has three serious motivations. First, and most obviously, there is the motive of applied ethics: sophisticated mind control methods may someday become reality, and we need to know what to say about them. Second, as a psychiatrist, I am interested in the many criticisms leveled at my field—one of which is that psychiatrists often wrongly control others' minds—and I hope that thinking about very extreme sorts of mind control may shed light on those controversies. Finally, there is the motive of philosophical ethics, and in specific the matters of examining the ethics of consent and evaluating one common interpretation of the Kantian notion that we should respect others' rational personhood. This latter principle seems to provide the most natural account of why mind control would be wrong, so thinking about mind control may help us examine it.

In Section 1, I start by defining the term *mind control*, and in the process say a little about what I take rationality to be. I then briefly review the Kantian account, outline how it could explain the wrongness of things like brain rays, and compare it to similar positions taken about other sorts of manipulation to which persons may be subjected, such as lying. In Section 2, I point out that, despite the seeming impermissibility of mind control, we permissibly control others' minds all the time, in ways that are hard to differentiate from clearly impermissible sorts of mind control given only the resources of the Kantian view, and which cast that view into doubt. In Section 3, I consider some possible rejoinders to the examples used in Section 2 and the criticism of the Kantian view that follows from them. Finally, in Section 4, I offer the beginnings of an alternative theory, suggesting that a loosely contractualist approach may succeed in explaining both when mind control is permissible and when it is impermissible.

1. Background and Definitions

Throughout what follows, I use the term *mind control* to refer to what we might also call *anti-rational influence*, where this is causing someone to think or do something irrespective of whether she has adequate reason to think or do it, in a fashion that bypasses or undermines her rational capacities. By *think something* I mean *form or abandon a judgment-sensitive attitude* (Hieronymi 2006), where judgment-sensitive attitudes include beliefs, desires, and intentions.

The notion of rationality assumed in this paper is merely a formal one. That is, I want to assume only that being rational requires thinking and acting in the ways one has adequate reason to think and act, without taking a stance on what those reasons are. Put in terms of judgment-sensitive attitudes, we might say that being rational is a matter of adopting, through the exercise of one's rational capacities, those beliefs, desires, and intentions one has adequate reason to adopt. Being irrational is, on this view, a matter of failing to think or act in the ways one has adequate reason to act.

In what follows, I will not offer an account of the nature of reasons, taking this to be a ground-level notion. I do, though, want to observe that any plausible notion of rationality needs to assume that reasons are usually internal (Williams 1981), since we typically do not want to say that a person has been irrational just because her judgment-sensitive attitudes fail to reflect reasons which were not available to her anyway—ignorance and irrationality are different matters. Still, our notion of rationality also needs to incorporate standards of sensitivity to external reasons. Rationality is not merely a matter of thinking and acting according to the reasons one *takes* there to be, but is

also a matter of responding to reasons that there are. Although it would not necessarily make one irrational to have a false belief if the error is due to limited information, it might make one irrational to have a false belief if the error occurred because one was not appropriately attuned to evidence that contradicted it.

I want to avoid giving an account of what reasons we have—that is, to avoid giving a substantive account of rationality—for two interrelated reasons. First, which theory of practical rationality is correct is controversial: one could claim, for example, that to be rational is to maximize one's own utility or happiness, or instead that satisfaction of one's actual or even merely perceived interests is rational. Second, for present purposes, no opinion about what is substantively rational is necessary, as the anti-rational influences we will examine can be assumed to interfere in the *processes* of reasoning as well as in their practical consequences. Moreover, the Kantian view about rationality which will be my target is itself relatively silent about substantive matters.

Although generally anti-rational influences cause someone to think or act in ways she does not have adequate reason to think or act, they come in two different types.¹ First, there are *anti-rational influences properly considered*, where a person is influenced, by a mechanism that bypasses or undermines her rational capacities, to do something she would not have done, or chosen to do, otherwise—as when X uses a brain ray to cause Y to give X all of his money, but Y never would have thought to do so without that influence. Second are what we might term *para-rational influences*. In these cases, a person is influenced to do something she would have done anyway, but the influence still undermines or bypasses her rational capacities. For instance, an insurance salesman might motivate me to buy life insurance by frightening me with ghastly tales of what could happen to my family without it, where that fear undermines my rationality, even though I would also have bought the life insurance if I were given the opportunity to reason about the decision. In that I suppose it is harder to defend the permissibility of anti-rational influences properly considered, I will focus on the former, and assume that what holds for the anti-rational properly considered also holds for the merely para-rational.

Finally, it is important to point out that anti-rational influences, even properly considered, vary with respect to how dramatically they shift a person's attitudes and actions. One might influence someone to do something she would not have done otherwise but still induce only a small shift in her inclinations, as when Y is undecided between A and B, though leaning slightly more strongly toward B, but X persuades him to do A by giving him a mind control serum that makes him suggestible and then telling

1. Thanks to Jesse Summers for pointing out this distinction.

him to do A. Para-rational influences can also vary in this way, as when Y is more inclined to do B than A, but X nevertheless cements his decision by giving him the mind control serum.

Again, I assume that most of us regard very extreme, science-fictional sorts of mind control as quite clearly wrong. Joel Feinberg artfully expressed this sentiment when, in surveying the moral valence of all varieties of human oppression, he wrote:

[M]ost odious... is the manipulation of a person without his consent. Patients or prisoners ... can be drugged, put under total anesthesia, and then made to undergo lobotomies or other kinds of surgical manipulation or mutilation of the brain. Psychotropic drugs used in small quantities and electric stimulation of the brain for short periods have less severe effects and are revocable, but when imposed on a person without his consent ... they are hardly distinguishable on moral grounds from assault and battery. (Feinberg 1987, 67)

Feinberg himself is not much of a Kantian, but I suspect that to most modern philosophers the most natural account of the wrongness of things like brain rays and other technologically advanced types of mind control is basically a Kantian one. Kant famously claimed that we must always treat others' humanity never merely as a means, but also as an end in itself (Kant 2002). For Kantians, *humanity* is often equated with *rational personhood*, and Kant's admonition is understood to mean that we should never manipulate, undermine, corrupt, or exploit others' rational faculties. One might think that this is what makes lying wrong—by lying, I exploit your rational faculties to get you to do something—and it also seems to explain why it would be wrong to use a brain ray to force a person to think or act a certain way: in so doing, one would make him think something in an anti-rational fashion, thereby undermining or exploiting his rational faculties and showing disrespect for his humanity. For Kantians, this is not only a wrong, but the most fundamental of wrongs.

Though I am not aware of anyone who has made this particular argument about mind control in the philosophical literature, related positions abound. For instance, Ginger Hoffman has recently argued that it is morally wrong to take antidepressants without engaging in psychotherapy because it *objectifies* oneself—that is, runs roughshod over one's rational personhood—by making one think things without engaging one's reason (Hoffman 2013). The Kantian stance also underpins much of the literature on coercion and undue influence, and is widely regarded as the basis for the doctrine of informed consent in bioethics (Beauchamp and Childress 2001). There, it is commonly supposed that what

makes consent essential is our obligation to respect another person's rational autonomy (his capacity to govern his own life through the exercise of his rational capacities), and it is only when an agent lacks our usual rational capacities (because he is developmentally disabled or mentally ill or already under the control of someone else) that we need not respect his decisions about his life.

2. Permissible Mind Control

Accordingly, I take it that the Kantian account of the wrongness of mind control is natural and appealing, even if it has not been widely defended. I also think it has some worrisome inadequacies, at least in the relatively simple form presented above. In particular, we can show that rationality need not always be respected: ordinary life abounds with cases where it is *permissible* to make others believe or do things without giving them good reasons, thereby undermining or circumventing their rationality. For example, consider the following case:

Connie Consumer wants to buy a watch. She is in a department store, contemplating both a cheap digital watch and a Rolex. Suppose she is more inclined, initially, to buy the cheap digital watch. Sally Salesperson thinks Connie is better off with the Rolex; she can tell by the way that Connie is dressed that she is wealthy, and it is a better watch for the money. So Sally persuades Connie to buy the Rolex. She does this in part by offering Connie reasons for buying it. But she does other things, too: she puts the Rolex under a bright light on a black background, so it sparkles appealingly. She draws Connie's attention to a nearby advertisement depicting a beautiful woman strolling down the Champs Elysée wearing the same watch. She models the Rolex for Connie, handling it reverently, but treats the digital watch as though it were slightly soiled. Ultimately, Connie buys the Rolex.

I assume that what Sally does in this case is permissible: she is just engaging in ordinary sales tactics, and these are a widely tolerated practice about which we have few, if any, moral qualms. But I also think some of Sally's actions amount to anti-rational influence, in that she gets Connie to (want to) do something without offering her reasons to (want to) do it. It seems plausible, for instance, that Connie is influenced, at least marginally, by the poster displayed behind the display counter. But when Connie decides to buy the Rolex, she does not think: "oh, that's a really nice poster there, with that beautiful model wearing this watch. *That's* a reason for me to buy it." Or, if she did think anything like

that, she would be irrational (or at least mistaken), since she would be taking herself to have a reason she does not. Rather, the poster presumably increases the likelihood that Connie will buy the Rolex by causing her to associate the watch with romance, exotic settings, and other things that she desires—even though having the Rolex typically would not make the satisfaction of those other desires more likely.

Likewise, the facts that the Rolex sparkles appealingly beneath the display lights and looks good on Sally's wrist are not reasons for Connie to buy it. They may highlight some reasons to buy it, namely the facts that the watch is beautiful and that it would look attractive on her own wrist. But Connie is probably already aware of these reasons. If placing the Rolex under the lights and modeling it affect her decision, they do so either by causing her to pick the Rolex without deliberation, or else changes the salience such reasons have for her, perhaps by making her attend to them more fully.² Either way, these influences seem to subvert her rationality, by making her decision depend in part on factors outside of reason.

Finally, the different ways Sally treats the Rolex versus the cheap digital watch might affect Connie's decision in an anti-rational fashion. Obviously, one reason to buy a Rolex instead of a cheap digital watch is that Rolexes have a certain cachet: they increase your social status, at least marginally. Sally's ability to influence Connie's decision by handling the watches differently plays upon this reason, but presumably is neither itself a reason (that *Sally* will think more highly of Connie if Connie buys the Rolex is generally not, absent some special story, a good reason for Connie to buy it), nor even reminds Connie of the reasons that she has (phenomenologically, it seems implausible that Connie recognizes what Sally's different treatments of the watches signify about their effects on status, and then chooses the Rolex on the basis of its potency as a status-symbol). Rather, as anyone who has been influenced in this way may be able to attest, it is something like the triggering of her instinctive desire to achieve social status or to display its trappings that allows Sally to influence Connie in this manner.

Sally's ability to influence Connie is just one of many examples furnished by modern consumer marketing, where permissible mind control techniques abound. Obviously, print, television, and internet advertisers have caught onto the fact that putting scantily-

2. It is well known that even the physical proximity of a desired object can influence persons' decision-making. For instance, Bushong et al. (2010) showed that subjects' willingness to pay for snack-food items (a measure of how much they want those items) is significantly higher if those items are physically present rather than simply displayed on a computer screen, even if they will be delivered in the same time-frame. It is also widely accepted that drug addicts are more likely to relapse if they see the drug of abuse or are given other reminders of their use (Robbins and Everitt 1999).

clad persons in their advertisements can improve the sales of almost anything. This strategy may be unethical because it exploits the models, but the fact that it slightly compromises consumers' rationality does not seem to make it so. Celebrity endorsements have the same effect (Spry et al. 2011). The fact that a popular NASCAR driver appears on television endorsing motor oil he is paid to endorse is no reason whatsoever for me to buy that motor oil, but may influence me even so. Again, however, this influence seems morally permissible. Physical stores use anti-rational influences, too. In addition to the techniques of salespersons, grocery stores will do things like locating bread, cheese, milk, and meat near the rear of a store, so we have to wander long aisles full of things we were not initially planning to buy to make our planned purchases; likewise, many groceries contain in-house bakeries that make small quantities of fresh bread throughout the day, often at a loss, because the smell of fresh bread is comforting and makes people hungry, inclining them to buy more in the rest of the store (*Economist* 2008).

Nor are anti-rational influences restricted to the commercial world. They occur in other aspects of social life, too. It is clear our physical appearance can affect others' thoughts about us in subtle ways. We all know that our relative pulchritude can alter how others treat us—how likely they are to bestow favor upon us, how they assess the merits of our ideas, and how much authority they cede to us. The psychological literature bears this out (Baker and Churchill 1977). Likewise, we frequently take pains to exploit this fact: we might wear a fancy suit to a job interview, or submit an attractive photograph with a job application, with the clear aim of increasing our chances of success. I take it that, on the whole, these efforts are permissible, but I also suppose their effect is often anti-rational. While one might argue that personal hygiene and grooming are moderately reliable indicators of the qualities that employers usually desire, the attractiveness of a persons' face is probably not (unless the employee is being hired for a position where attractiveness is itself essential).

For a somewhat different example, note that I sometimes wear glasses when I am giving lectures or presentations. I could wear contact lenses instead. But suppose I were to choose to wear glasses on some occasions because I know from having read the psychological literature that they subtly but measurably improve my audiences' implicit assessments of my intelligence, thereby (presumably) increasing the likelihood they will conclude my arguments are sound (Thornton 1994; Manz and Lueck 1978). I think most of us would regard my wearing glasses for this reason to be morally permissible. But again, it probably represents an anti-rational influence: my glasses do not provide anyone with a *reason* to conclude my arguments are sound: people who wear glasses are not more likely to make sound arguments.

Some things we do to influence others in anti-rational ways quite closely resemble habits and reflexes and may be hard to distinguish from them. Facial expressions, posture, and gestures all affect others' thoughts about us and their responses to what we say and do, but generally do so without giving reasons. As a psychiatrist, I sometimes try to look more sad and worried for my patients than I may actually feel in the moment. Although I invariably care about them and want them to get well, occasionally it is hard to muster genuine empathy for them; one's emotions can become fatigued. Fortunately, after some practice, I know how to make myself look and sound like I feel sad and dismayed even when I do not feel that way. I also know that getting my patients to believe I am sad or dismayed can make a difference to their treatment: it can make them more likely to reveal important information and more likely to follow my recommendations about their care. Admittedly, it could be somewhat discomfiting to them to learn I do this, but I do not suppose the practice is unethical. On the contrary, mustering affective displays that are somewhat insincere seems required both by my concern for my patients' well-being, as well as by my professional (and ultimately, moral) obligation to treat them as well as I am able. But once again, it seems clear that influencing my patients through these artificial displays of emotion represents an anti-rational influence: I manage, I suppose, to make them have beliefs (as well as feelings) about me that make them more likely to do things I think they ought to do, without giving them genuine reasons to do those things (though clearly I assume they have other reasons to do those things).

In sum, then, I take it that these examples strongly suggest that anti-rational influences—including both para-rational influences and anti-rational influences properly considered—can be morally permissible. But if that is the case, it follows that it is sometimes permissible to intentionally compromise others' rationality, and the Kantian's claim that we have an obligation to respect others' rational personhood, where this is taken to entail that we must refrain from doing things that compromise their rationality, must be false. This leaves us unable to explain why mind control techniques that seem obviously impermissible are so.

3. Objections and Responses

Of course, proponents of the Kantian position have a variety of responses to the foregoing argument at their disposal, which it should be worthwhile to explore here.

First, one might simply insist that all the sorts of "permissible" mind control described above are really impermissible. This insistence is not unmotivated. After all, it has been argued before that advertising with a persuasive (as opposed to merely informative)

focus is impermissible (Santilli 1983; Crisp 1987) and one can imagine taking the stance that advertising in general—prone as it is to making us want things we do not need, leading to increased dissatisfaction or, alternatively, to run-away consumption, the waste of precious resources, environmental degradation, and so forth—ought to be prohibited (Carlson et al. 1985). This is not exactly a Kantian motivation, but we can also concede that advertising looks more distasteful the more closely we scrutinize it, that certain sorts of advertising techniques (e.g., subliminal advertising) are clearly impermissible, and that the more effective advertising is in motivating a person to purchase an item—especially when the influence is properly-considered anti-rational—the less permissible it seems.

Even admitting these other considerations, however, I think we should not follow the Kantians in drawing this hard line. Though the matter comes close to being a fundamental disagreement, we should note two things. First, the Kantian position would commit us to a revisionist stance regarding a large number of practices that now appear to be morally permissible—such as dressing nicely to improve our prospects of getting a job, wearing makeup, mirroring a patient’s emotions in a therapy session, or even ordinary sales tactics. This seems to me to be enough reason to reject the Kantian view. Second, we should note that this is really just one species in a larger genus of problems for Kantians, so that a revisionist stance with respect to these cases would commit us to revisionist stances elsewhere. In particular, the standard Kantian admonition to respect rational personhood is also taken to apply to our *own* rational personhood, and would seem to exclude the use of alcohol and other mind-altering substances that many of us regard as at least occasionally permissible. Conversely, if we are willing to make exceptions in this sort of case, then we should be more willing to admit exceptions in cases of mind control.

One might also object that the plausibility of the cases above hinges on an equivocation between anti-rational influences properly speaking and para-rational influences. According to this objection, the examples of mind control above seem permissible largely insofar as they are para-rational, but para-rational influences need not be regarded by Kantians as impermissible. The motivation for this objection is that para-rational influences in general seem less problematic than anti-rational influences properly speaking. For instance, it is clearly permissible for a commercial to be used to motivate me to purchase a new car if I am already inclined to purchase a new car (and have sufficient reasons to do so), but it is less clearly permissible for someone to use a commercial to motivate me to buy a car when I had no such inclination antecedently or, worse, was inclined *against* buying a car.

Still, this objection rests on two mistakes. First, it is false that all of the cases above are correctly interpreted as (merely) para-rational and not as anti-rational properly speaking. In the case of Connie Consumer, we assume at the outset that Connie is inclined to buy

the cheap digital watch instead of the Rolex. Nor is this assumption implausible. More generally, the net effect of things like consumer marketing techniques is not, overall, to get people to do things they were going to do anyway, but to do things they were not going to do otherwise (it would otherwise not be economically efficient). One supposes that sales techniques can increase a consumer's willingness to pay, the sales of a particular item, or the overall consumption of a particular commodity—all of which would be best explained by a capacity to influence others to do things they would not have done in the absence of the influence.

Second, it is not at all clear that Kantians should be friendlier toward para-rational influences than they are toward properly-speaking anti-rational influences (that is why I have persisted in placing them in the same category, anti-rational influences generally). The Kantians' position looks to the mechanism of an influence, rather than its outcome—what matters is not *what* an agent decides to do as the result of our interactions with him, but *how* we brought him to decide to do it, and different influences may undermine a person's rationality to the same extent even if one is para-rational and the other properly anti-rational. Indeed, the Kantian requires a process-oriented as opposed to outcome-oriented position about what makes mind control wrong in order to make sense of certain intuitions, such as that it would be wrong to use a brain ray to manipulate a person into doing something he would choose to do anyway. The Kantian takes the same position about coercion: that it is wrong to coerce someone into doing something even if he has a reason to do it and even if he would decide to do it on his own, because coercion subverts his rational personhood and thus his autonomy.

So this second objection seems to fail. There are, however, other ways to reinterpret the cases. For instance, many Kantians accept that we can make it permissible for others to do things to us that would usually be impermissible if we consent to that treatment in advance. Often, implicit or tacit consent is adequate for this purpose. Thus, one might claim that what makes Sally's manipulation of Connie, or advertisers' manipulation of consumers generally, or my influence on my patients—is that in each case the parties influenced know that they are going to be influenced before proceeding, and by proceeding give tacit consent to that influence. The plausibility of this objection is suggested by a revision of Connie's case: imagine that instead of acting like the salesperson she is, Sally masquerades as another shopper but still does many of the things she did in the original example to influence Connie's decision to buy the Rolex. Her role here seems more deceptive, and correspondingly more worrisome—presumably because Connie *did not* implicitly consent to being influenced in this way.

While we can concede that consent has a transformative moral force, this objection only passes muster if we assume that persons who enter a store, or watch television advertisements, or who come into a psychiatrist's office for a consultation, *know* that they are likely to be subjected to the influences outlined above. But that seems implausible, both because this does not seem to be the sort of thing of which people are typically cognizant, and because many of these influences would, one supposes, be rendered ineffective if we were aware of them. This certainly seems true of my patients' responses to my emotional displays, and (though I am aware of no psychological data that supports this claim) seems likely to be true of sales techniques and advertising.

For a fourth objection, one might be tempted to explain the influences in the cases above as permissible because they ultimately *enhance* our rationality or are parasitic on processes that are essentially rational. For instance, suppose Sally partly brought about Connie's decision to buy the Rolex just by taking it out of the display case, increasing its physical proximity to Connie. One could claim that this influence depends on a rational (albeit implicit) judgment on Connie's part, one that reflects discounted utility theory. People tend to value rewards that are more temporally proximate over equivalent, and even larger, rewards that are temporally distant (Ainslie 1975). Discounted utility theory claims that this is rational: we *should* discount rewards at a rate that is proportionate to their temporal distance, since the more distant in time a reward is, the less likely it is that we should ever receive it. Similarly, one might claim that Connie's adjusting her relative valuation of the Rolex when it became more physically proximate reflects an underlying decision-making mechanism that embodies a rational principle, according to which rewards are discounted more the farther away from us they are.

This objection also fails, however. First, even assuming it is sometimes rational to discount utility by physical proximity, this seems like a spurious rationalization for Connie's decision, since in this particular case the physical proximity of the Rolex says little about the likelihood that Connie will be able to obtain it (that depends mostly on how large her credit card limit is). More generally, showing that our decisions sometimes invoke a mechanism that is rational as a general strategy or heuristic or which may have been evolutionarily advantageous does not make those decisions rational themselves. Finally, the literature on temporal discounting suggests that we do not always respect discounted utility theory: decision-makers do not apply their discounts in a consistent fashion, exhibiting what is known as dynamic inconsistency. We will often, for instance, value reward A over the larger reward B when A is immediate and B is temporally removed by time T1, but will value B over A when A is delayed by time T2 and B is delayed by T1 + T2, so that we have discounted B by proportionately more than A in the first case but

not the second (Kirby and Herrnstein 1996). This violates discounted utility theory, and implies that the underlying structure of our decisions is, in this respect, irrational.

4. An alternative account

If you take any of the examples I gave above as cases of permissible mind control, then one has to conclude the Kantian position I outlined earlier is inadequate, because it is sometimes permissible to do things to others that undermine or exploit their rational faculties. Now I want to offer an alternative to the Kantian position. The alternative has three elements, and aims both to explain why things like brain rays would be wrong and why mundane sorts of mind control are often permissible.

First, let me define what I'll call the relative *violence* represented by a particular mind control technique. Violence, in this context, is a composite of two other properties, namely *power* and *invasiveness*. Power is the propensity of a technique, considered broadly across an array of possible settings, to induce persons to form judgment-sensitive attitudes in a way that does not reflect the balance of reasons. Wearing glasses during a lecture is not a particularly powerful influence: it is unlikely to persuade you that I am making good arguments if I am not, for example. In contrast, a brain ray or a well-told lie can be quite powerful, in that it is (by hypothesis) able to get people to form attitudes contrary to the reasons they have.

Invasiveness is related to power but importantly different. It is a technique's capacity to affect aspects of a person's mind or thinking that are more or less central to his identity, irrespective of the balance of reasons. Suppose there is a brain ray that can affect persons' attitudes only when the balance of reasons is equivocal, so it is not powerful in the sense outlined above. Now assume that I am a Zoroastrian and someone uses the brain ray to turn me into a Buddhist. Even supposing I had no reasons to accept Zoroastrianism over Buddhism, the forced change in my religious beliefs seems extremely objectionable—and thus, violent in my sense of that term. Presumably, what makes it violent is that it changed a characteristic essential to my identity.³

3. This example raises an interesting question: since most religious conversions presumably do not hinge on the balance of reasons, but almost always affect core components of identity, shouldn't it nearly always be impermissible to try to convert someone? Most of us are inclined to think that it is not, but that raises a question about whether invasiveness really matters at all. I am tempted to say that standard religious conversions actually are not invasive in the sense I intended because, even though they change features of a person's beliefs that are central to his identity, they do so in an identity-preserving fashion: that is, they depend upon the exercise of his rational faculties and typically this results in a set of beliefs with which the agent actively *identifies*. But this response fails, since we can imagine a brain ray that would work in more or less the same manner, and which

My second observation is that the permissibility of any particular mind control technique reflects, to some degree, whether its relative violence is sufficiently counterbalanced by the expected risks and benefits to all concerned parties, compared to alternative courses of action. Overall, it seems like even extremely violent mind control techniques can permissibly be used in extreme circumstances, and that even relatively non-violent techniques should not be used if the expected harms are large and the expected gains are small. Together, these factors account for some of the following intuitions:

- a. It is often wrong to use especially violent mind control techniques even if the aim is beneficent. For example, I ought not use telepathy to make a patient take his antihypertensive medications, even if it would be better for him, and even if no other methods of persuasion exist.
- b. It can be permissible to use even very violent techniques of mind control if the benefit/harm balance is sufficiently large. For instance, Professor X can permissibly use his telepathic powers to stop Alice from shooting someone or crashing her car into a wall.
- c. It is often impermissible to use even non-violent mind-control techniques when the expected harms are very large and not outweighed by the associated benefits: for instance, it would be wrong for Marilyn Monroe to use her feminine wiles to persuade President Kennedy to give her nuclear access codes so she could sell them for a profit.
- d. Finally, others can permissibly treat us in anti-rational ways to benefit themselves, even if there are costs to us, as long as the benefits to them tend to outweigh our costs and the treatment is not too violent; examples include when the grocery store owner or advertiser or car salesman manipulates us into buying things we might be better off without or into spending more money than we might otherwise do, or an interviewee's charming smile persuades us to choose him instead of a more qualified candidate.

would still seem impermissible.

A different factor we should consider is that in the standard cases, a person is (I assume) usually converted by a true believer. This seems to be important to the permissibility of the conversion, in that it would be morally very different to convert someone to a set of religious beliefs one does not accept oneself (whether that be for one's amusement or for material gain). In forming our intuitions about conversion-by-brain-ray cases, we may be assuming that the brain-ray operator does not accept the beliefs to which he converts others.

But in the end, I think a contractualist account is essential to differentiating these cases; for more on that, see below.

I should emphasize, however, that I do not suppose—or at least am not myself able to identify—anything like a formula for balancing the violence of a mind control technique with its risks and benefits. Nor do I even suppose that the loosely consequentialist stance outlined above amounts to the whole story. It is merely one set of considerations among potentially many others that inform our judgments about what is permissible.

This brings us to the third element of my proposal. I think we can understand the way we balance risks, benefits, and violence when determining the permissibility of mind control in a roughly contractualist way. I do not have the space to delve into the details of contractualism here, nor to defend contractualism as a moral theory. It should be sufficient to remark that according to most contractualist moral theories, actions are impermissible if and only if it would be reasonable for us to choose, under unbiased conditions, to prohibit them (Scanlon 1996). So consider clearly impermissible sorts of mind control first. Plausibly, we could all reasonably agree, if we were impartial, not to subject each other to extremely violent methods of mind control, at least unless the situation were dire. This is particularly obvious when such methods would impose severe harms on us to only slightly benefit others—mind control in those cases would be wrong for the same reasons that theft and violence often are. Moreover, it is reasonable for us to be disposed against especially violent mind control techniques even when the net benefit of their use in a particular case is great, because the temptation to utilize such techniques inappropriately is high: it behooves us to assume there is a standing prohibition against such influences even if that prohibition is not absolute. Something similar is true of coercion: although it can be permissible to coerce others when the balance of risks and benefits is clearly favorable (as when we do so to prevent violence against others or to encourage them to adopt behaviors, like wearing seatbelts, that have great benefit but minimal cost), we do well to use coercion only with great caution.

On the other hand, given a contractualist background, we can explain why certain mundane, non-violent mind control techniques are permissible. This is because they represent, or are at least closely *related to*, essential aspects of human social practices which could not feasibly be eliminated from our lives, either because they are so often beneficial to us collectively, or because any efforts to prohibit them in some cases would alter other aspects of our lives undesirably. Given such features, it would not be reasonable to choose together to prohibit them *even though* they compromise our rationality. We want to be allowed to smile at people even when we do not feel happy, to dress nicely on important occasions even if this may be misleading, and possibly, to wear glasses when we give philosophy lectures, because these types of interventions make social life better, or are at least practically indistinguishable from other actions that do. And, I suspect, the value

we attach to our ability to do these things is great enough that we are willing to tolerate others' occasionally performing similar actions in ways that undermine our rationality. Even if, in a more ideal world, we could interact with each other on a purely rational level, our capacities to influence others' thinking in anti-rational ways, and perhaps even our capacity to be so influenced, are desirable features of our embodied personhood. If the alternative to being occasionally gamed and tricked and confounded by others is adopting a hypersensitivity to anti-rational influences that would dramatically alter the sorts of creatures that we are, we should clearly choose the former.

The contractualist reasoning offered above obviously will not hold for all sorts of permissible anti-rational influence. For instance, it does not seem to apply to persuasive advertising techniques, which are not, so far as I know, integral to human social life. But I suspect that in many of these cases other contractualist accounts are available. In the case of advertising, one might suppose that it is rational for us to allow persuasive advertising despite its potential for anti-rational influence for something like the following reasons: (1) most of us greatly value having access to print, television, radio, and online media; (2) the production and distribution of these media are costly and in general those costs cannot be covered by subscription fees; (3) persuasive advertising that is itself visually captivating or entertaining but which also has the potential for influencing consumers in anti-rational ways is an efficient method for covering media costs; (4) alternative sorts of advertising that are purely informative and accordingly bland and uninteresting, but which also would not influence anyone in anti-rational ways, are not efficient ways of covering media costs (because they are less effective than persuasive advertising at capturing consumers' attention and affecting their purchases); (5) in the main, the extra costs imposed on us by advertising are sufficiently counterbalanced by the benefits of access to media; and, (6) persons who do not wish to be influenced by persuasive advertising can simply limit their exposure to the media that contain it. Of course, whether these considerations are sufficient to render persuasive advertising permissible depends on whether each of considerations (1)-(6) above is true, which is mainly an empirical matter. My purpose here is not to make an argument for the permissibility of advertising *per se*, but just to illustrate the sorts of considerations such an argument might include.

Applying a loosely contractualist way of thinking to mind control has useful implications. For instance, it is a way of making sense of—and partially⁴ justifying—the

4. Only partially, however, because it is clearly possible that some of our moral reactions to novel technology are simply unfounded or irrational.

tendency many of us have to regard technologies as impermissible just because they are new and unfamiliar. For most of us it would seem *more* problematic to increase others' estimations of our intelligence by using a brain ray than it would be to increase their estimations of our intelligence by wearing glasses, even if we assume the power and the invasiveness of the two techniques are equivalent (i.e., that this is all the brain ray in question can do). This seeming discrepancy can sometimes make sense from a contractualist viewpoint, for two reasons. First, wearing glasses to impress others with our intellectual ability bears a significant resemblance to other things we reasonably want to be free to do, such that it might not be reasonable for us to prohibit the practice. This does not hold for using a brain ray to induce similar beliefs. Second, at least given the way the case above is described, we have little information about the different ways the brain ray technology in question could be developed or applied, and thus little information about the other sorts of manipulation it could make possible. Although to some extent we could assess each of these future developments on its own merits as it occurred, we also recognize that social discourse about the permissibility of novel practices (not to mention actual legislation) often lags far behind practices themselves (this is perhaps one of the main motivations for bioethics: that we should think about the moral implications of new technologies before we use them, because it is often hard to go back). Thus, permitting the use of a novel technology might seem unreasonable, even if the potential harms of the technology are not significantly different than those of practices with which we are already familiar, just because we cannot reliably, either theoretically or in practice, differentiate that technology from others it would clearly be reasonable to prohibit.

5. Conclusions

The foregoing arguments suggested that the Kantian contention that our rational nature should be inviolable is inconsistent with how we usually live our lives, wherein we permissibly utilize all manner of anti-rational influences. This implied that Kantians are unable to account for why some mind control techniques are impermissible, paving the way for an alternative theory. To be sure, there is nevertheless much that is correct about the Kantian view—our rationality is undeniably a good, and a distinctively human good at that. Indeed, one wonders whether the best way of accommodating Kant's recommendation to respect others' rational personhood does not involve something along contractualist lines. Contractualism has its roots in Kant's ethics, and many contractualists (such as Scanlon) think that our obligation to abide by rules that we would reasonably choose together is underpinned by an obligation to respect each other as rational. It

would of course, be ironic if it turned out that respect for rationality sometimes permitted undermining rationality, but this seems consistent with much of our experience. In the end we are not *ideally* rational beings, and we are not *merely* rational beings: much of what we do, and much of what we think, is the product of irrational, or at least non-rational, aspects of our selves. Although our rationality is valuable, it is also sometimes *rational* to trade it against other considerations that have value; we are complicated and messy and emotional, and it is sometimes best to embrace these facts about ourselves.

References

- Ainslie George. 1975. "Specious reward: a behavioral theory of impulsiveness and impulse control." *Psychological Bulletin* 82 (4): 463–496.
- Baker, Michael and Gilbert A. Churchill Jr. 1977. "The impact of physically attractive models on advertising evaluations." *Journal of Marketing Research* 14 (4): 538–555.
- Beauchamp, Thomas L. & James F. Childress. 2001. *Principles of biomedical ethics*. New York: Oxford University Press.
- Bushong, Benjamin, Lindsay M. King, Colin F. Camerer, and Antonio Rangel. 2010. "Pavlovian processes in consumer choice: The physical presence of a good increases willingness-to-pay." *The American Economic Review* 100 (4): 1556–1571.
- Carson, Thomas L., Richard E. Wokutch, and James E. Cox Jr. 1985. "An ethical analysis of deception in advertising." *Journal of Business Ethics* 4 (2): 93–104.
- Crisp, Roger. 1987. "Persuasive advertising, autonomy, and the creation of desire." *Journal of Business Ethics* 6 (5): 413–418.
- Feinberg, J. 1987. *Harm to Others. The Moral Limits of the Criminal Law. Volume I*. New York: Oxford University Press.
- Hart, Joshua, James A. Schwabach, and Sheldon Solomon. 2010. "Going for broke: Mortality salience increases risky decision making on the Iowa gambling task." *British Journal of Social Psychology* 49 (2): 425–432.
- Hieronymi, Pamela. 2006. "Controlling attitudes." *Pacific Philosophical Quarterly* 87 (1): 45–74.
- Hoffman, Ginger A. 2013. "Treating Yourself as an Object: Self-Objectification and the Ethical Dimensions of Antidepressant Use." *Neuroethics* 6 (1): 1–14.
- Kant, Immanuel. 2002. *Groundwork for the Metaphysics of Morals*. New Haven: Yale University Press.
- Kirby, Kris N., and Richard J. Herrnstein. 1995. "Preference reversals due to myopic discounting of delayed reward." *Psychological Science* 6 (2): 83–89.
- Manz, Wolfgang, and Helmut E. Lueck. 1968. "Influence of wearing glasses on personality ratings: Crosscultural validation of an old experiment." *Perceptual and Motor Skills* 27 (3): 704–704.
- Robbins, Trevor W., and Barry J. Everitt. 1999. "Drug addiction: bad habits add up." *Nature* 398 (6728): 567–570.
- Santilli, Paul C. 1983. "The informative and persuasive functions of advertising: A moral appraisal." *Journal of Business Ethics* 2 (1): 27–33.
- Scanlon, Thomas M. 1998. *What we owe to each other*. Cambridge: Harvard University Press.

- Spry, Amanda, Ravi Pappu, and T. Bettina Cornwell. 2011. "Celebrity endorsement, brand credibility and brand equity." *European Journal of Marketing* 45 (6): 882–909.
- "The science of shopping: the way the brain buys." *Economist* December 18, 2008.
Retrieved from: <http://www.economist.com/node/12792420>
- Thornton, George Russell. 1944. "The effect of wearing glasses upon judgments of personality traits of persons seen briefly." *Journal of Applied Psychology* 28 (3): 203.
- Williams, Bernard. 1981. "Internal and external reasons." In *Moral Luck*. Cambridge: Cambridge University Press, 101–13.