PROJECT MUSE®

Access Provided by Fondren Library, Rice University at 01/08/13 11:21PM GMT

J. S. Blumenthal-Barby

Between Reason and Coercion: Ethically Permissible Influence in Health Care and Health Policy Contexts

ABSTRACT. In bioethics, the predominant categorization of various types of influence has been a tripartite classification of rational persuasion (meaning influence by reason and argument), coercion (meaning influence by irresistible threats—or on a few accounts, offers), and manipulation (meaning everything in between). The standard ethical analysis in bioethics has been that rational persuasion is always permissible, and coercion is almost always impermissible save a few cases such as imminent threat to self or others. However, many forms of influence fall into the broad middle terrain—and this terrain is in desperate need of conceptual refining and ethical analysis in light of recent interest in using principles from behavioral science to influence health decisions and behaviors. This paper aims to address the neglected space between rational persuasion and coercion in bioethics. First, I argue for conceptual revisions that include removing the "manipulation" label and relabeling this space "nonargumentative]influence," with two subtypes: "reason-bypassing" and "reason-countering." Second, I argue that bioethicists have made the mistake of relying heavily on the conceptual categories themselves for normative work and instead should assess the ethical permissibility of a particular instance of influence by asking several key ethical questions, which I elucidate, that relate to (1) the impact of the form of influence on autonomy and (2) the relationship between the influencer and the influenced. Finally, I apply my analysis to two examples of nonargumentative influence in health care and health policy: (1) governmental agencies such as the Food and Drug Administration (FDA) trying to influence the public to be healthier using nonargumentative measures such as vivid images on cigarette packages to make more salient the negative effects of smoking, and (2) a physician framing a surgery in terms of survival rates instead of mortality rates to influence her patient to consent to the surgery.

Kennedy Institute of Ethics Journal Vol. 22, No. 4, 345–366 © 2012 by The Johns Hopkins University Press

INTRODUCTION: BACKGROUND, MOTIVATION, AND PLAN FOR PAPER

In bioethics, the predominant categorization of various types of influence has been a tripartite classification of *rational persuasion* (meaning influence by reason and argument), *coercion* (meaning influence by irresistible threats—or on a few accounts, offers), and *manipulation* (meaning everything in between; Faden, Beauchamp, and King 1986). The standard ethical analysis in bioethics has been that rational persuasion is always permissible, and coercion is almost always impermissible. However, many forms of influence fall into the broad middle terrain—and this terrain is in desperate need of conceptual refining and ethical analysis.

This is especially true given the recent interest in using behavioral science insights to change individual and group health related decisions and behaviors through techniques that fall somewhere in between reason and coercion. Examples include but are not limited to the use of *subconscious cues* to trigger healthy behaviors, *incentives* for weight loss and smoking cessation, *framing and focusing effects* to get patients to focus on certain risks or benefits, *default* HIV screening for all adults entering hospitals (recommended by the Centers for Disease Control), default Sickle Cell Trait screening for all college athletes (recommended by the National Collegiate Athletic Association), and *vivid images* on cigarette packages (now mandated by the FDA).

The use of behavioral economics and behavioral psychology principles to shape health decisions and behaviors is likely to grow due to several factors. In the United States, the National Institutes of Health (NIH) has made "The Science of Behavior Change" a priority by designating it as a Roadmap Initiative (National Institues of Health 2009); the NIH and the Agency for Healthcare Research and Quality awarded 7 grants totaling 20 million dollars on "Behavioral Economics for Nudging the Implementation of Comparative Effectiveness Research"; the U.S. Department of Agriculture awarded 2 million dollars in 2010 for research on how behavioral economics can improve federal food policy; and the Robert Wood Johnson Foundation awarded eight $100,000 grants in 2012 to study "Applying Behavioral Economics to Perplexing Health and Health Care Challenges" such as obesity and lack of consumer engagement. In the U.K., the Institute for Government and the Cabinet Office published a ninety-six-page report, "MINDSPACE: Influencing Behaviour Through Public Policy," exploring how behavior change theory can meet policy challenges; the Department of Health issued "guidance on the most effective behaviour

change techniques" in December 2010 in its Draft Structural Reform Plan; and The House of Lords Science and Technology Committee launched an inquiry into the effectiveness of behavior change interventions, which it released in July 2011. Given all of this, attention to the space between reason and coercion is essential.

Thus, this paper aims to add some conceptual clarity to and ethical analysis of this space. It focuses on forms of influence that work by relying on facts about people's psychology such as knowledge about their emotions, how they perceive things, how they make judgments and decisions, and what they desire. First, I argue for conceptual revisions that include relabeling much of this space "nonargumentative influence," and viewing the category as having two main subcategories, "reason-bypassing nonargumentative influence" and "reason-countering nonargumentative influence." Second, I argue it has been a mistake to rely heavily on conceptual categories themselves for normative work (e.g., X is an instance of coercion and is therefore morally problematic, or X is an instance of manipulation and is therefore morally problematic). Instead, I argue that in order to assess the ethical permissibility of a particular instance of influence in health care and health policy contexts, we should ask several key ethical questions that relate to two main ethically relevant dimensions regarding influence: (1) the impact of the form of influence on autonomy and (2) the relationship between the influencer and the influenced. Regarding the autonomy dimension I argue that, contrary to the standard view, nonargumentative influence does not *necessarily* interfere with autonomy, or validity of consent. To assess whether it does, the following key questions should be asked: (a) Are options significantly blocked or burdened? (b) Is the person aware of the fact that they are being intentionally influenced and of the mechanism of that influence? (c) What is/would be the person's attitude towards the fact that they are being influenced by mechanism X in context Y? Regarding the relationship dimension, I argue that the nature of the relationship goes a long way towards determining the moral valence of nonargumentative influence. To assess the ethical permissibility, the following key issues should be explored: (a) what the influenced could reasonably expect and view as appropriate in the context of that relationship, (b) what obligations the influencer has in the context of that relationship (e.g., a high degree of transparency, promotion and fostering of autonomy, etc.), and (c) whether the instance of influence would damage the relationship by implying lack of respect, equality, ability, etc.

THE PREDOMINANT ACCOUNT: RATIONAL PERSUASION, MANIPULATION, AND COERCION

As mentioned in the introduction, the predominant typology of influence used in bioethics is the one developed by Ruth Faden, Tom Beauchamp, and Nancy King. They categorize the types of influence into three types: rational persuasion, manipulation, and coercion. On their account, rational persuasion is influence by the use of reason and argument, coercion is influence by the use of irresistible threats of harm, and manipulation is nonrational but noncoercive influence by altering choices available or altering a person's perception of choices available (Faden, Beauchamp, and King 1986, p. 261). In this predominant typology, much falls into the category of manipulation. Examples of manipulation include influence by incentivizing, offering, increasing options, decreasing options, tricking, using [resistible] threats of punishment, managing information, presenting information in a way that leads to predictable inferences, deceiving, lying, withholding information, slanting information, and exaggerating it in a misleading way (Faden, Beauchamp, and King 1986). Manipulation can also involve misleading packaging or misleading images, trading on fear, subliminal suggestion, flattery, guilt, appealing to emotional weakness, and initiating psychological processes that are difficult to reverse or that lead to predictable behaviors or decisions.

PROBLEMS WITH THE PREDOMINANT ACCOUNT

There are two main problems with Faden, Beauchamp, and King's typology. The first is with the labeling of the category "manipulation." On their account, many cases get labeled as manipulation that most of us would not consider manipulation under ordinary usage of the term. For example, on their account, when a professor tells her students that if they attend a lecture they will get extra credit, she is manipulating them. When a spouse makes flattering remarks to get his partner to cheer up at the end of a tough day he is manipulating her. Insofar as we want our typology to map onto ordinary usage, this one does not since too many acts or behaviors get labeled as cases of manipulation that do not strike most people as cases of manipulation given the usual negative connotations of the term. The related yet bigger problem with the account is with the breadth of the category of "manipulation." This becomes especially problematic when moving to the normative stage where one is expected to make claims about the moral status of "manipulation." I will expand on this point below in my suggested conceptual revisions.

SUGGESTED CONCEPTUAL REVISIONS

I suggest that bioethics will be better suited to consider the ethics of influence if we do two things:

(1) Classify forms of influence based on the following schema (see Figure 1):

(a) Reason and Argument
(b) Nonargumentative Influence (Reason-Bypassing Type)
(c) Nonargumentative Influence (Reason-Countering Type)
(d) Omission
(e) Force or Severe Threats

(2) Cease to rely heavily on the categories themselves for normative work.

Let me comment on the second point first. It may be the case that the conceptual categories give us hints about the ethical issues that should be considered, but they do not themselves do substantive normative work. For instance, as I will argue in the next section, there will be examples of nonargumentative influence that pose a threat to autonomy and examples that do not. Whether it does depends on the context, persons involved, etc. and not on it being a case of nonargumentative influence per se. In short, bioethicists have been mistaken to approach the question of influence as if we can develop some sort of definitive judgment about "*the* moral status of nonargumentative influence," or on the old account "*the* moral status of manipulation."

And let me now comment on my suggested classification schema. First, on the categories themselves, I think it is fairly clear what Influence by Reason and Argument, Influence by Omission (of information or options), and Influence by Force or Severe Threats mean. By Reason-Bypassing Nonargumentative Influence I mean influence that operates by bypassing a person's reasoning capacities and often their awareness, with examples including framing, setting up defaults, setting up the environment a certain way, and priming using subconscious cues. And by Reason-Countering Nonargumentative Influence I mean influence that operates by countering a person's reasoning capacities, with examples including social norms/pressures, inducing affective states, playing on desires. I recognize that influences that play into a person's affect or desires do not necessarily lead her to something that runs counter to what her reasoning capacities led her to; however, it is typical both in philosophical analyses and in ordinary usage to draw a distinction between reason and emotion or desire, with the idea that emotion and desire often counter or operate as

an antithesis to a person's reasoned decisions. At any rate, the distinctive feature of Reason-Countering Nonargumentative Influence is that, unlike Reason-Bypassing Nonargumentative Influence, a person is often quite aware that the influence is occurring and that it operates by inducing or playing into desires or emotions.

Second, these five categories (or their subtypes) are not mutually exclusive. For example, an offer of money to participate in a weight-loss program might fall into both categories of Influence by Reason and Argument and Reason-Countering Nonargumentative Influence. An offer is certainly something that a person considers as she engages in a reasoning process of weighing the pros and cons of a course of action, but it is also something that moves the person by exciting her desires. Also, this typology is dependent on context. Context will sometimes affect where an instance of influence falls in the typology. For example, if a person is being influenced by a social norm in a way that she is not aware of (e.g., she is told that 86% of the population wears their seat belts and so she is subconsciously primed to put on her seat belt) then it is an example of Reason-Bypassing Nonargumentative Influence. On the other hand, if a person is influenced by a social norm in a way that she is aware of (e.g., her boss designates a certain day as "wear your 'I got my flu shot' sticker to work day" and she feels pressured to get a flu shot) then it is an example of Reason-Countering Nonargumentative Influence.

I am sure that this typology is not complete, nor is it neatly divided. Yet, it is an improvement on the existing rational persuasion–manipulation–coercion typology in several ways. First, it offers more detail for the territory between reason (influence by reason and argument) and coercion (influence by force or severe threats). Second, it avoids the morally laden term "manipulation" and instead uses more descriptive terminology, which allows us to separate the categorization of a type of influence and the moral analysis of it.

ETHICAL ANALYSIS

The ethical analysis of a particular instance of influence depends, I will argue, on two components: (1) the impact of this form of influence on autonomy and (2) the relationship between the influencer and the influenced. Before I proceed, I want to make two clarifying points about my focus. First, I focus my analysis on the space between reason and coercion that is nonargumentative influence since it is the space that has received the least amount of attention in bioethical analysis, and it is the space that needs

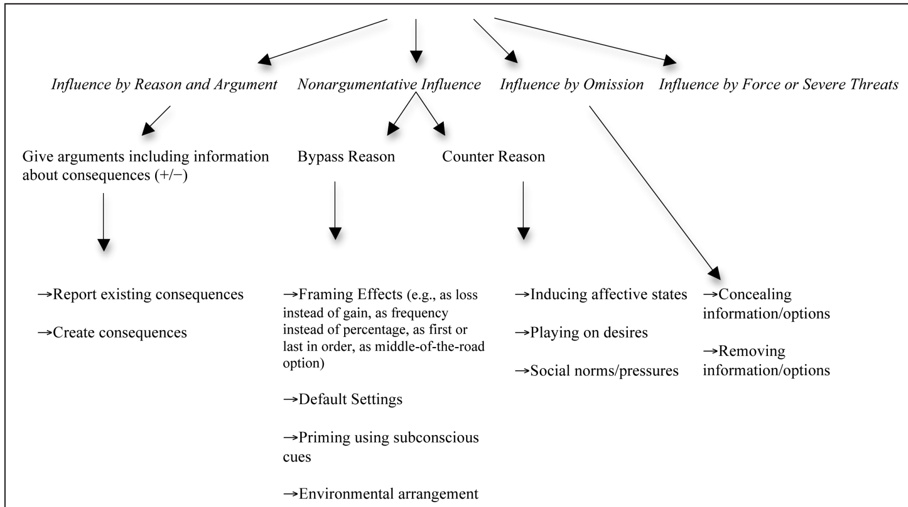

Figure 1. Types of Influence

the most attention in light of recent attempts to influence people's health behaviors and decisions through the use of behavioral science insights that operate in these ways. Second, I leave out any explicit discussion of the *end* towards which an influence attempt is directed. While that is certainly a component of the ethical analysis of a particular instance of influence, the controversy exists in the use of nonargumentative influence *even in* cases of good ends. Moreover, there are of course tangential epistemological and ethical issues about how we decide what is good for an individual or society, and how we balance individual goods and societal goods, but those are beyond the scope of this paper.[1] In this paper I focus on ethical components beyond the usual focus on good ends. Those two additional factors are impact on autonomy of the influenced and nature of the relationship between the influenced and the influencer.

AUTONOMY

The Standard View: Incompatibility of Nonargumentative Influence and Autonomy

The standard ethical analysis, in every account that I know of, is that influence by reason and argument (so long as reasons used are believed by the influencer to be true) is morally unproblematic. In fact, giving someone reasons and arguments for why they should do one thing or another shows respect for them as an agent, moral or otherwise. Influence by force

or threats of harm, on the other hand, is generally thought to be morally problematic, and hence justified only under a narrow set of circumstances, such as when a person poses an imminent risk of harm to himself or others or violates societal laws. Influence by nonargumentative means has settled somewhere in between these ends, but I think that it is fair to say that it has engendered a good deal of moral suspicion, and I think that it is fair to say that much of this comes from concerns about autonomy.

Autonomy involves shaping one's own life in ways that one finds valuable or important, as opposed to going through life mindlessly or based on other people's agenda (Dworkin 1988, p. 164). Thus, so some argue, influence that bypasses or works against a person's reasoning abilities (abilities that allow him to consider how he wants to govern himself and what sorts of things he wants to pursue or avoid or aspire to) poses a threat to self-governance or autonomy. Melissa Seymour Fahmy (2011, p. 183) has objected that such influence usurps the agent's authority to direct her life as she sees fit, Robert Noggle (1996, p. 52) has argued that it thwarts the agent's rational and moral agency, and John Martin Fisher (2004, p. 13) has argued that free will requires that a person's behavior is appropriately sensitive to *reasons*. Stanely Benn (1967, p. 135), Gerald Dworkin, (1988, p. 18), Patricia Greenspan (2003, p. 159), and Marcia Baron (2003, p. 50) have made similar points. Three final examples of the standard view that nonargumentative influence and autonomy are incompatible: Eric Cave argues that in nonargumentative influence A mobilizes a "nonconcern motive" to get B to behave differently than B originally would have and as a result A is managing B's concerns, not B, and as such B's autonomy is violated (Cave 2006, pp. 138–139). Similarly, in his recent book *Practical Autonomy and Bioethics*, James Stacey Taylor has argued that the absence of "manipulation," or nonargumentative influence, is a necessary condition for autonomous action. Taylor asserts

> It is clear that the successful manipulation of a person into performing an action that she would not have otherwise performed would serve to compromise the manipulee's autonomy with respect to her manipulated actions. (Taylor 2009, p. 41)[2]

Finally, Daniel Hausman and Brynn Welch argue that all forms of influence other than rational persuasion, including these nonargumentative influences, or "nudges," interfere with autonomy by diminishing the extent to which the agent has control over her evaluations and deliberations (Hausman and Welch 2010, pp. 128, 135).

Problems with the Standard View: Why Nonargumentative Influence and Autonomy Can Be Compatible

Despite the predominance of this narrative that nonargumentative forms of influence pose a threat to autonomy, I will argue that it is mistaken. While I support the general notion that in order for people to govern themselves and manage their concerns, they have to engage their reflective and reasoning capacities, it is too quick to claim that every instance of influence that bypasses or counters these capacities usurps autonomy. There are several authors who have gone against the grain and taken the position that nonargumentative influence and autonomy can be compatible, specifically, Faden, Beauchamp, and King; and Sarah Buss and Harry Frankfurt. I will explicate their arguments, pointing out problems where they occur, and then offer revisions in my account of the conditions under which the compatibility holds.

Interestingly, Faden, Beauchamp, and King have taken the position that influence that operates by going around or against reason does not always pose a threat to autonomy. Commenting on the moral justifiability of manipulation (and recall by this Faden, Beauchamp, and King just mean any nonargumentative influence not rising to the level of irresistible threats) they write, ". . . some manipulative influences are controlling, whereas others are compatible with parts of the influence continuum of influence that permit autonomous action" (1986, p. 354). Whether autonomy is compromised depends on (1) the ease or difficulty the manipulatee finds in attempting to resist, and (2) the extent to which the influence interferes with the manipulatee's substantial understanding (Faden, Beauchamp, and King 1986, p. 365).

Faden, Beauchamp, and King's account is problematic in two ways. First, the resistibility criterion for measuring when an instance of influence compromises autonomy overlooks a large portion of cases that we are concerned about, i.e., any case where influence works by *bypassing* an agent's reasoning (e.g., cases where someone conceals an option or a piece of information, cases where someone uses subconscious cues). In those cases, the ethical concern is not that the influence was so strong that the person could not resist it; it is that the person was not even aware that she was being influenced. Second, the substantial understanding criterion for measuring whether an instance of influence compromises autonomy will include too many cases (i.e., every case of decision or action based on inadequate information).[3] Requiring that understanding is necessary for autonomous choice and action is controversial for this reason. As James

Stacey Taylor convincingly argues with the example of Martin Fobisher—the English explorer who thought he was making trips to collect gold, but was really collecting "fool's gold"—who was certainly extremely self-governed in those excursions despite the fact that he did not understand what he was bringing back (Taylor 2009, p. 6).

Second, Sarah Buss has gone against the grain and argued that the wrong (insofar as there is one) in influence via nonargumentative forms of influence such as "manipulation" (which she defines as influencing preferences and beliefs nonrationally)[4] is *not* to be found in violations of autonomy or self-governance. She argues that (1) being influenced via nonargumentative forms of influence is entirely compatible with self-governance, and (2) many well-informed, self-governing agents would endorse a policy that involved being influenced by nonargumentative forms of influence. To support the first point, she argues that we consider people to be self-governing or autonomous despite the fact that there is a past and an external world that influences their current mental states and actions, and that nonargumentative influences are no different. If the past and the external world do not pose threats to self-governance or autonomy than neither do nonargumentative influences from others (Buss 2005, p. 212). Buss makes the same argument about ignorance. We consider people to be self-governing or autonomous despite the fact that they are ignorant about all sorts of aspects of their situations and choices (this is like Taylor's point with the Fobisher fool's gold example), so being ignorant about the fact that they are doing something because of a nonargumentative influence is no different (Buss 2005, p. 214). Moreover, Buss argues, we are constantly influenced by nonrational means, and it cannot possibly be that in all of these cases people are not able to self-govern or act autonomously (Buss 2005). Incidentally, Harry Frankfurt has made almost this exact same point. Frankfurt writes, "We are inevitably fashioned and sustained, after all, by circumstances over which we have no control. . . . It is irrelevant whether those causes are operating by virtue of the natural forces that shape our environment or whether they operate through the deliberatively manipulative design of other human agents" (Frankfurt 2002, p. 28).[5]

To support her second point—that many well-informed, self-governing agents would endorse a policy that involved being influenced by nonargumentative forms of influence (presumably leading to the conclusion that the influence is autonomy-preserving based on many procedural accounts of autonomy such as that of John Chrisman (2009))—Buss argues that

we should simply remind ourselves of the wide range of circumstances where nonargumentative influence occurs and where we are not the least bit threatened. In fact, we often welcome or value it, such as when we are seduced by a lover (Buss 2005, p. 218, 220).

Overall, Buss makes a compelling argument for the compatibility of nonargumentative influence and autonomy, but there are a couple of weaknesses. First, on the Buss–Frankfurt point that it makes no difference for autonomy whether the nonargumentative influence comes from a random environment or from the intentional design of another person, I am ambivalent. There is surely a metaphysical difference between the two, and perhaps a moral distinction, but is there a distinction that is important for autonomy? Consider an example where I am making a decision of whether or not to undergo a surgical procedure, yet I am underinformed about the risks simply because it is a new procedure and that information is not yet available. Contrast that with an example where I am making a decision of whether or not to undergo a surgical procedure, yet I am underinformed about the risks because my doctor has intentionally kept that information from me in order to get me to undergo the surgery. It seems to me that in the latter case the physician has imposed his will upon mine in a way that presumably I would not endorse and that poses a threat to my autonomy or self-governance in a way that simply not having the information does not. The reason that "other" as cause vs. environment as cause may be relevant for autonomy is that an agent is more likely to repudiate a desire/decision formation that is directed by the will of *another* than they are a random one.

This relates to a second weakness in Buss's account. Contrary to her claim that when we reflect on nonargumentative influences in our decision procedures we would mostly endorse them, it is equally easy to imagine that upon reflecting on how often we are subject to nonrational influence, we would chastise ourselves for falling prey to it and vow—from now on—to rely more on our critical perspective and reasoning skills. It is also easy to imagine that even though a person might not object to the fact that she is often not as rational or reflective as she ideally might be, she may indeed object to another person taking advantage of that fact to lead her in one direction or another. And finally, just because many well-informed, self-governing agents would endorse nonrational influence in romantic contexts, it does not mean that they would endorse nonrational influence in other contexts (e.g., in a health care context). I may have no objections to being influenced by my husband's charisma to get me to take out the

garbage, but I may have objections to being influenced by my physician's charisma to take a medication.

When Autonomy and Nonargumentative Influence Are Compatible: Key Normative Questions

Some cases of influence pose a threat to autonomy and some do not. The key normative questions with respect to autonomy that we should be asking ourselves about various instances of influence are these:

1. Are options significantly blocked or burdened?
2. Is the person aware of the fact that she is being intentionally influenced and of the mechanism of that influence?
3. What is/would be the person's attitude towards the fact that she is being influenced by mechanism X in context Y?

Question 1 gets at concerns about liberty, or a person's ability to do as he wishes, and Questions 2 and 3 get at concerns about autonomy, or a person's ability to shape his own life or be intentionally shaped by others in ways that he is aware of and approves of, or at least does not repudiate. It is important to remember, normatively speaking, that while an instance of influence posing a threat to liberty or autonomy does have a wrong-making feature that speaks as a reason against it, this does not mean it is ipso facto morally wrong. *First,* an *instance* of influence may usurp someone's autonomy at that time, but promote or protect her autonomy in the long run and as such be morally justifiable even from an autonomy-grounded moral theory. Consider a competent patient who is hesitant to take a psychiatric medication and a physician who knows that if the patient does not take the medication then in the future he will lose the ability to make important medical decisions. In response, the physician uses a form of nonargumentative influence that had the patient known about it he would likely not endorse (e.g., presenting the risks of the medication as sandwiched between description of the benefits) and as such would infringe on his autonomy at that moment, but for the sake of preserving and promoting his capacity to exercise his autonomy in the future. *Second,* there are other considerations besides for autonomy that enter the moral picture. For example, public health officials may choose to implement forms of nonargumentative influence that would infringe on the autonomy of some individuals for the sake of a substantive public good (e.g., fines for polluting). Such examples raise issues about balancing individual goods (including autonomy) and societal goods, but again, in

this paper I set those issues aside to focus on ethical components beyond the usual focus on ends. Having addressed the impact on autonomy of the influenced, I now turn to the second major factor to consider in the ethical analysis of the permissibility of nonargumentative influence in health care and health policy contexts: the nature of the relationship between the influenced and the influencer.

RELATIONSHIP BETWEEN INFLUENCER AND INFLUENCED

Obligations and Expectations Between Parties

The second major factor at play in the ethical permissibility of nonargumentative influence is the relationship between the two parties and the duties that relationship implies. The nature of a relationship goes a long way towards determining the moral valence of nonargumentative influence. Consider four examples. In Example One a storeowner arranges a display to draw attention to the expensive items and distract attention from the cheap items, influencing her customers to purchase the expensive items. In Example Two a wife, knowing her husband's love of tools, purchases a new lawn mower, influencing her husband to mow the lawn.[6] In Example Three a physician, knowing that people are more influenced by relative risk information than absolute, informs a high-risk patient of her breast cancer risk as compared to the risk of most women, influencing the patient to take the preventative drug tamoxifen. In Example Four a governmental institution begins a campaign to get people to eat healthier, taking advantage of the power of subconscious cues and priming by pumping the smell of delicious fruits through the vents of public transportation systems.

I would venture that most of us would not object to the storeowner case and the spouse case as being morally problematic. On the other hand, the physician case and the government case may raise objections. Part of the reason is because we *expect* advertisers and romantic partners to influence us by nonargumentative means as a matter of routine interaction, whereas we do not expect our physicians and our governments to operate in these ways. Thus, an important question to ask is: *What does the influenced reasonably expect from the potential influencer?* It is important because reasonable beliefs related to trust generate moral obligations that form the basis for judgments of about ethical permissibility.

In the physician–patient relationship, for example, we expect certain things, i.e., a high level of transparency, reasons for the actions the physician is recommending, promotion of our best interests, and support and promotion of our autonomy (which is more than it just not being violated).

These are appropriate and reasonable expectations for that relationship, and the physician who uses the nonargumentative influence of relative risk framing to get her patient to take tamoxifen is not supporting or promoting the patient's autonomy. The physician may actually be violating the patient's autonomy insofar as the patient is not aware of the fact that she is being intentionally influenced and had she been aware that she was being influenced by that framing mechanism by her doctor she would not endorse that as a process by which she was satisfied with making the decision to take tamoxifen. But, even if the physician is not violating the patient's autonomy (imagine the patient would not repudiate having made her decision because of the physician's influence mechanism), the physician is not promoting the patient's autonomy. She is not encouraging the patient to consider the available options in light of her preferences, values and goals and come to a decision that is best in line with those.

We can contrast this with contexts where promoting autonomy is not important, such as the advertising context. There may even be health contexts where promoting autonomy is not as important, as in the context of directly observed therapy for the treatment of tuberculosis (health care providers watch patients swallow every dose) where health care providers working for public health departments have a stronger obligation to protect the public. We consider it a reasonable and appropriate expectation that our governmental health care agencies protect the public from epidemics, and acknowledge that that sometimes involves restricting the liberty of individuals.

Damage to Relationships

Because this paper focuses on contexts of health care and health policy, I want to note an aspect of the relationship factor that goes beyond the obligations of one party to the other, namely that in health care, the relationship between physician and patient is importantly built on trust. Insofar as nonargumentative influence damages the physician–patient relationship, it is ethically problematic. Nonargumentative influence might damage the relationship in any of these cases:

(1) if the patient feels that in failing to engage her reasoning capacities the physician is failing to show respect for her by treating her as lesser, not an equal, and not capable;

(2) if the patient feels that in failing to engage her reasoning capacities the physician is failing to show respect for her by outright dismissing her views and judgments as not worthwhile; or

(3) if the patient feels that the physician is exploiting her weaknesses (Baron 2003, p. 50).

Relationship Between Influencer and Influenced: Key Normative Questions

Thus, just as it was important to ask certain questions about particular instances (type and context) of influence regarding autonomy (e.g., whether the instance of influence significantly blocks or burdens options, whether the person is aware that she is being intentionally influenced and of the mechanism of that influence, and what is/would be the person's attitude towards her being influenced by a particular mechanism in a particular context), it is important to ask questions about particular instances of influence regarding the relationship between the influencer and the influenced:

1. What could the influenced reasonably expect and view as appropriate in the context of that relationship?
2. What obligations does the influencer have in the context of that relationship including whether they have the obligation to promote and foster autonomy?
3. Would the instance of influence damage the relationship by implying a lack of respect, equality, ability, etc.?

Here are two contrasting examples that get at these points well. Doctor 1 thinks that Patient 1 needs a surgery. Doctor 1 shows Patient 1 a vivid video of what might happen if the patient does not get the surgery. Doctor 1 does this with all of the humility in the world, thinking that the video is the best way to get the patient to deeply appreciate the consequences in a way that sitting down and talking through the reasons with the patient would not. Patient 1 senses the care and motivation with which the doctor presents the video and perceives it as such. Doctor 2 also thinks that Patient 1 needs surgery and Doctor 2 also shows Patient 1 a vivid video of what might happen if the patient does not get the surgery, but Doctor 2 shows the patient the video because he thinks that Patient 1 is an idiot who would not understand or appreciate the reasons that he would give the patient were they to sit down and talk through it, and he thinks that Patient 1 is weak and overly emotional and thus the video would trigger the patient's emotions and fears. Patient 1 senses this and perceives the video as a scare tactic, damaging the relationship. In both cases we have the exact same nonargumentative influence mechanism, but Case 2 is ethically problematic because of damage posed to the physician–patient relationship.

APPLICATION OF THE ETHICAL ANALYSIS

Application to a Public Policy Case

I now turn to examining how the analysis that I have provided would apply to two real-life cases of "manipulation," or behavioral influence by nonargumentative means. One existing case that has generated controversy is the case of governmental agencies such as the Food and Drug Administration (FDA) trying to influence the public to be healthier using nonargumentative measures such as vivid images on cigarette packages to make more salient the negative effects of smoking. I will address first the extent to which this practice promotes or threatens autonomy by looking at the extent to which it blocks or burdens options, and the extent to which the affected members of the public are aware of and endorse (or were they aware of would endorse) the use of vivid images by the government as the process by which their decision not to smoke or buy cigarettes was formed.

We face an initial epistemic challenge posed by this criterion because it is difficult to know for certain whether subjects would or would not endorse the process. Of course, we can make some intelligent guesses, and we can begin to empirically investigate whether the public would endorse this process of desire/behavior formation or not. But there is a larger problem: insofar as we are concerned about autonomy, we are concerned about the autonomy of *individuals*. It makes little sense to talk about the autonomy of the public. Therefore, there is a sense in which it is not helpful to make guesses about or empirical investigations about what processes of desire/behavior formation most people would endorse. For example, although it is reasonable to guess that, and it may turn out to be the case that, most people would endorse governmental use of vivid images on cigarette packages to increase the salience of the negative consequences of smoking, there are inevitably going to be several people who would absolutely repudiate the government's involvement in the process of their desire/behavior formations. Perhaps the best that we can do in public policy cases with respect to autonomy is to aim to preserve the autonomy of the *majority* of individuals, such that the autonomy-related question at the level of public policy is whether the majority of people would endorse the nonargumentative influence. And my intuition in this particular case is that the majority of people would endorse, or at least not repudiate, that process by which their decision not to buy a pack of cigarettes or smoke was formed.

And now, to what extent does the FDA's use of vivid images to make the negative effects of smoking salient to the public fulfill or violate expecta-

tions and obligations that arise in that relationship, and preserve or damage the relationship? The FDA is a governmental body employed to protect and promote public health. As such, they do indeed have an obligation to protect and promote health, and it is reasonable for the public to expect that they would work to decrease tobacco use, given that it is an addictive drug that results in many negative health consequences. In some ways no risks of damage to the relationship are posed by the campaign. After all, the FDA does not fail to engage the public with arguments and reasoning because they view the public as incompetent, but rather because (1) to truly engage in dialectic reasoning about the pros and cons of smoking would be highly impractical if not impossible, and (2) they are countering the nonargumentative influence employed by cigarette advertisers. The efforts of the FDA are aimed at respecting the worth of individuals by taking up concern for the health of all citizens, not just those of wealthy upper socio-economic status. It is well documented that the majority of smokers are of a lower socio-economic class. On the other hand, the public may perceive the campaign as consisting of scare tactics that demonstrate an intimidating and power-abusing federal agency. Hausman and Welch make a similar point when they raise the concern that “nudges” from the government may be seen as disrespectful towards citizens (Hausman and Welch 2010, p. 138). For this reason empirical work can and should be done to test public perception before a technique is employed. Moreover, the public and key stakeholders can and should be engaged in the design and implementation of such interventions or policies so that they feel a part of the project and relationships are preserved as opposed to damaged.

Application to an Interpersonal Case

The second case that I analyze differs from the first because it is not a public policy case, but a case occurring at the individual–individual level. It also differs from the first which is an example of nonargumentative influence that operated by using facts about subjects’ psychology (e.g., knowledge that they are influenced by what they find salient, and motivated by fear) in an obvious way; this second case involves nonargumentative influence that operates by bypassing subjects’ awareness. Case two is this: a physician frames a surgery in terms of survival rates instead of mortality rates to influence her patient to consent to the surgery.

The extent to which this practice promotes or threatens the autonomy of the patient depends on the extent to which the patient is aware of and endorses (or would endorse were she aware of) this framing by the

physician as part of the process by which her decision to consent was made. Whether or not the process would be endorsed likely depends on how it is described. If the process is described as follows, the patient may not be inclined to endorse it: "Your decision to consent was formed by a process in which your physician framed things in a certain way (i.e., survival rates instead of mortality rates) to increase the likelihood that you would consent to surgery." If, however, the process is described as follows, the patient may be more inclined to endorse it: "Your decision to consent was formed by a process in which your physician reflected on whether to frame things in terms of chances of survival or mortality and decided to frame things in terms of survival because she did not want to frighten you and thought the surgery was best in line with your values and goals," or "because she knows that you are likely to overestimate the risks and thought the surgery was best in life with your values and goals."

It is also interesting to note that the extent to which an instance of nonargumentative influence promotes or threatens autonomy may depend on the way in which, and the level of detail with which, we describe it, at least insofar as we are operating under historical accounts of autonomy such as Christman's. Further, note that the patient might endorse part of the process of nonargumentative influence (e.g., the physician nonargumentatively influencing), but not another part (e.g., the particular nonargumentative influence mechanism of framing). This problem is really a meta-problem with historical accounts of autonomy such as Christman's, but insofar as it is arguably the predominant and least problematic account of personal autonomy, we are pressed to work this out. One approach is to focus on whether the influenced individual is *on the whole* satisfied with that nonargumentative process of decision/behavior formation.

An astute observer might note that it seems as if autonomy concerns can be reduced to concerns about whether an instance of nonargumentative influence is reasonable given that answers to questions of endorsement will likely turn on the reasonableness of the attempt. Thus, the observer might argue, our concern is not about autonomy at all, but only about whether the influence attempt is well intended, and is reasonable or expected in the particular context in which it is employed. In fact, Buss has made this exact argument in her paper on manipulation (Buss 2005, p. 218). While I believe that the questions about endorsement and reasonableness are closely linked, I think that it is a mistake to collapse the concepts. Autonomy and reasonableness are separate concepts and concerns. Moreover, the autonomy question and the endorsement question will not always

have the same answer. The patient may see it as perfectly reasonable that a physician would try to influence a patient to make healthy decisions, but still not endorse her decision being formed in that way.

To finish the application of the analysis to the example, let us now examine the extent to which framing the surgery in terms of survival rates in order to get the patient to consent fulfills or violates expectations and obligations that arise in that relationship, and preserves or damages the relationship. As noted earlier, the relationship between a physician and a patient involves certain obligations and expectations. Physicians have an obligation to protect and promote the health interests of their patients, but they also have an obligation to protect and promote their autonomy. Patients may reasonably expect that their physician will try to protect and promote their health, but they also reasonably expect that their physician will be straightforward with them. I do not expect my mother to be straightforward with me when she tries to promote my health interests, I expect her to exaggerate about the negative effects of smoking for example, but I do expect my physician to be. To make due on her expectation to be as straightforward as possible, I would expect the physician to go ahead and initially frame the surgery in terms of survival rates, but also say to me, "That also means that so and so number of people die from the surgery." The situation might be different when we are discussing my yearly failure to exercise, in which case I expect a little less straightforwardness and a little more nonargumentative influence. For example, I would expect her to remind me that 500 people die every year from not exercising, without also saying, "That also means that millions of people do not!"

Whether the framing to get the patient to consent to surgery will damage the relationship between the physician and the patient depends on whether the patient discovers the influence attempt and on how the patient responds to it. Whether discovery is damaging will depend on the individual psychology of the patient and the physician, and their individual relationship, and as such it is difficult to make general claims here. That said much hinges on whether the patient perceives the physician as acting from care and courage vs. arrogance and laziness. Moreover, as I have emphasized before, trust is an essential component of the physician–patient relationships, so the physician should carefully consider whether the patient would perceive this framing as a violating of trust if discovered.

CONCLUSION

In conclusion, we can have morally problematic influence and non-problematic influence. Whether or not an instance of influence is ethically permissible or not should not depend on what category it fits into per se (e.g., manipulation or coercion); instead, it should depend on the answer to certain key normative questions that I hope to have elucidated here. Consideration of these questions is important as we move to integrate the psychology of influence into medicine and health policy.

ACKNOWLEDGMENTS AND SUPPORT

A draft of this paper was presented at the Junior Scholars in Bioethics Workshop sponsored by The Center for Bioethics, Health and Society at Wake Forest University. My thanks to the conference organizer Ana Iltis, and to faculty who provided me with helpful comments on the paper: Eric Juengst, Rebecca Walker, Mark Hall, Christine Coughlin, and Ana Iltis. A thank-you is also due to my colleague, Baruch Brody, who had several helpful conversations with me about this paper, and to an anonymous reviewer from *KEIJ* who made excellent suggestions for improvement. Finally, I am grateful to the Greenwall Foundation for a Faculty Fellowship in Bioethics (2011–2014) and to the Pfizer Foundation for a MAP Fellowship in Bioethics (2011–2013), which support my work in this area.

NOTES

1. For a discussion of these problems see, for example, Blumenthal-Barby (2012).
2. Although Taylor does admit that this man could still be acting "of his own free will" even though he is not acting "autonomously" because for Taylor free will is about the way an individual's will is structured (i.e., about identification with one's desires, actions, etc.), whereas autonomy is a political concept about self-governance as opposed to other-governance (Taylor 2009, p. 41).
3. Of course, one might still hold that a lack of understanding poses a threat for the instrumental *value* of the exercise of autonomy in that one is less likely to exercise one's autonomy in a way that allows achieving goals or well-being if not well informed.
4. Buss is using the term "manipulation" simply to refer to nonargumentative forms of influence: ". . . the distinguishing mark of manipulative "processes" is that they influence preferences (and beliefs) nonrationally . . ." (Buss 2005, p. 208, n. 19). She is using the term "deception" to refer to ". . . the intentional misleading of one person by another . . ." (Buss 2005, p. 209, n. 19).

5. For Frankfurt, what matters is whether a person is wholeheartedly behind the desires that move him to act regardless of how he got them.
6. Example from (Andre 1985, p. 111).

REFERENCES

Andre, Judith. 1985. Power, Oppression and Gender. *Social Theory and Practice* 11 (1): 107–122.

Baron, Marcia. 2003. Manipulativeness. *Proceedings and Addresses of the American Philosophical Association* 77 (2) (November 1): 37–54. doi:10.2307/3219740.

Benn, Stanley. 1967. Freedom and Persuasion. [An earlier version of this paper was read to the Eighth Annual Conference of the Australasian Political Studies Association, Canberra, August, 1966.] *Australasian Journal of Philosophy* 45 (3) (January): 259–275. doi:10.1080/00048406712341211.

Blumenthal-Barby, J. S. 2013. Choice Architecture: A Mechanism for Improving Decisions While Preserving Liberty? In *Paternalism: Theory and Practice*, ed. Christian Coon and Michael Weber. Cambridge University Press.

Buss, Sarah. 2005. Valuing Autonomy and Respecting Persons: Manipulation, Seduction, and the Basis of Moral Constraints. *Ethics* 115 (2) (January): 195–235. doi:10.1086/426304.

Cave, Eric M. 2006. What's Wrong with Motive Manipulation? *Ethical Theory and Moral Practice* 10 (2) (December): 129–144. doi:10.1007/s10677-006-9052-4.

Christman, John. 2009. The Historical Conception of Autonomy. In *The Politics of Persons: Individual Autonomy and Socio-Historical Selves*, 133–163. New York: Cambridge University Press.

Dworkin, Gerald. 1988. *The Theory and Practice of Autonomy*. Cambridge University Press.

Faden, Ruth R., Tom L. Beauchamp, and Nancy M. P. King. 1986. *A History and Theory of Informed Consent*. New York: Oxford University Press.

Fahmy, Melissa Seymour. 2011. Love, Respect, and Interfering with Others. *Pacific Philosophical Quarterly* 92 (2) (June 1): 174–192. doi:10.1111/j.1468-0114.2011.01390.x.

Fischer, John Martin. 2004. Responsibility and Manipulation. *The Journal of Ethics* 8 (2) (January 1): 145–177.

Frankfurt, Harry. 2002. Frankfurt-Style Compatibilism: Reply to John Martin Fischer. In *Countours of Agency: Essays on Themes from Harry Frankfurt*, ed. Sarah Buss and Lee Overton. Massachusetts Insititue of Technology Press.

Greenspan, Patricia. 2003. The Problem with Manipulation. *American Philosophical Quarterly* 40 (2) (April 1): 155–164.

Hausman, Daniel M., and Brynn Welch. 2010. Debate: To Nudge or Not to Nudge. *Journal of Political Philosophy* 18 (1) (March): 123–136. doi:10.1111/j.1467-9760.2009.00351.x.

National Institues of Health. 2009. *NIH Science of Behavior Change*. Meeting Summary. Bethesda, Maryland.

Noggle, Robert. 1996. Manipulative Actions: A Conceptual and Moral Analysis. *American Philosophical Quarterly* 33 (1) (January 1): 43–55.

Taylor, James Stacey. 2009. *Practical Autonomy and Bioethics*. New York: Routledge.