

## **Framing Effects Do Not Undermine Consent<sup>1</sup>**

### **Abstract**

Suppose that a patient is receiving treatment options from her doctor. In one case, the doctor says, “the surgery has a 90% survival rate.” Now, suppose the doctor instead said, “the procedure has a 10% mortality rate.” Predictably, the patient is more likely to consent on the first description and more likely to dissent on the second. This is an example of a framing effect. A framing effect occurs when “the description of [logically-equivalent] options in terms of gains (positive frame) rather than losses (negative frame) elicits systematically different choices.” Framing effects are ubiquitous, but they are particularly troublesome in medicine. Many worry that there is tension between valuing informed consent and using framing effects in clinical settings. In this paper, I answer this question: if an individual is subject to a framing effect when she gives her consent, does this undermine the validity of her consent? I argue that framing effects do not undermine consent in general.

### **Keywords**

Consent; Framing Effects; Nudging; Trust

## **Framing Effects Do Not Undermine Consent**

Suppose that a patient is receiving treatment options from her doctor. In one case, the doctor says, “the surgery has a 90% survival rate.” Now, suppose the doctor instead said, “the procedure has a 10% mortality rate.” Predictably, the patient is more likely to consent on the first description and more likely to dissent on the second. This is an example of a framing effect. A

---

<sup>1</sup> Acknowledgements: thanks are due to David Boonin, Eric Chwang, Chris Heathwood, Iskra Fileva, Ben Hale, Mike Huemer, Frank Beckwith, Robyn Waller, Connor Kianpour, Bill Glod, Travis Reider, audience members at St. Norbert College, and audience members at the 2022 Eastern APA. Thanks are also due to Emily Director for all of her continued support.

framing effect occurs when “the description of [logically-equivalent] options in terms of gains (positive frame) rather than losses (negative frame) elicits systematically different choices” (Gonzalez et al., 2005: 1).<sup>2</sup> Framing effects are ubiquitous, but they are particularly troublesome in medicine. Many worry that there is tension between valuing informed consent and using framing effects in clinical settings. In this paper, I answer this question: if an individual is subject to a framing effect when she gives her consent, does this undermine the validity of her consent? I argue that framing effects do not undermine consent in general.

In section I, I outline background. In section II, I explain why framing effects pose a problem for informed consent and how philosophers have attempted to resolve this problem. In section III, I argue that framing effects do not, in general, undermine consent. In section IV, I argue that there is nothing special about medicine that would make framing effects undermine consent for patients even if they do not undermine consent in general.

### **I. Background:**

Framing effects range from trivial to extremely serious. For example, “consumers who are more likely to use credit than cash when the price difference between them is described as a discount (gain) for using cash, but are more likely to use cash when that same difference is described as a surcharge (loss) for using credit, are vulnerable to a framing effect” (Chwang 2015: 271). And, in a more dire case, “a patient who consents to a medical intervention when its prognosis is described as 90% chance of survival but who would dissent if it were instead described as 10% chance of mortality is subject to a framing effect” (Chwang 2015: 271).

Intentional framing effects are a subset of the general phenomenon of nudging. A nudge is “any influence that is expected to predictably alter a person’s behaviour without

---

<sup>2</sup> The term was introduced by Kahneman and Tversky.

(substantively) restricting her options” (Simkulet 2017: 536). Since the work of Thaler and Sunstein (2009), there has been immense interest in the use of nudging in many areas of public policy. In medicine, some see nudging as a solution to the long-standing problem of balancing informed consent and beneficence toward patients. As Simkulet says, “some philosophers have argued that during the process of obtaining informed consent, physicians should try to nudge their patients towards consenting to the option the physician believes best” (2017: 536). In effect, nudges offer the hope of being able to have our cake and eat it too; we can honor consent *and* ensure that patients choose the treatment option that is in their best medical interests. Given that the literature on nudging is vast and includes many kinds of nudges, I won’t address nudging in general. I focus just on framing effects.<sup>3</sup>

## II. The Problem and Proposed Solutions:

One might initially wonder why framing effects are a problem for consent. They don’t coerce the agent,<sup>4</sup> by definition they are not deceptive, and they don’t make the agent irrational or less competent. So, what is the problem? The problem is that there is something troubling about a patient making a potentially life-altering decision that she would not choose if it were worded differently. Indeed, framing effects do predictably affect patients’ decisions.<sup>5</sup> Consent is supposed to protect our autonomy and allow us to live our lives as we wish. But, if something as trivial as a framing effect makes us give or not give consent, it seems that consent is not playing a robust, autonomy protecting role.<sup>6</sup> Furthermore, there may be no frame-less or neutral way of

---

<sup>3</sup> For a helpful categorization of the different nudges, see Blumenthal-Barby (2021: 23-25).

<sup>4</sup> As they don’t involve a threat, framing effects are importantly different than what is called verbal coercion in the sexual ethics literature (see Conly 2004).

<sup>5</sup> See McNeil et al., 1982: 1261.

<sup>6</sup> When an agent is subject to a framing effect, he falls afoul of the principle of “description invariance,” which holds that “when two expressions obviously describe the same possible outcome, that outcome should be represented and evaluated in the same way under each description” (Fisher 2020: 8).

informing a patient. So long as the message involves a positive and negative component, it is subject to framing effects. Since virtually all treatments have some positive and negative component, all descriptions of treatment options are subject to framing effects. The problem seems unavoidable and potentially damaging to the doctrine of informed consent.

Three solutions have been advanced. *First*, Cohen has defended the view that framing effects need not undermine consent and that doctors should use framing effects to nudge their patients toward the medically better decision.<sup>7</sup> *Second*, Hanna has argued that framing effects do undermine consent and that this should lead us to doubt the importance of consent and use framing effects to nudge patients in the medically better direction. *Third*, Chwang has argued that both Cohen and Hanna are wrong in some way. Chwang's view is that framing effects do undermine consent in medicine, but there are ways of removing their effects on patients by debiasing. Thus, doctors should neither use framing effects nor nudge their patients. They should attempt to debias them and proceed as normal. The position I defend in this paper will present new arguments to in support of conclusion similar to Cohen's.

Before proceeding, I'll address why I reject Chwang's solution. Before addressing whether framing effects invalidate consent, it is important to first consider whether we can avoid the problem altogether by debiasing. I am extremely doubtful about the possibility of successfully debiasing patients. As Blumenthal-Barby says, "there is little evidence that educating people about the phenomena in chapter 1 [referring to decisional heuristics and biases] works to rid them of them" (Blumenthal-Barby 2021: 89). She further argues that "because there are so many framing effects and biases that are triggered by so many small and nuanced aspects of the choice architecture, it is difficult to avoid them all, since things must framed one way or

---

<sup>7</sup> Cohen's (2013) defense is worded in terms of nudging in general.

another” (Blumenthal-Barby 2021: 89). For example, we could say both frames to the patient, meaning that we tell the patient the survival rate frame and the death rate frame. This seems initially plausible, because it puts all of the information on the table. However, this falls prey to the problem once again, because a framing effect, or more precisely an ordering effect, is introduced depending on which frame the doctor says first.<sup>8</sup> As Hanna puts it, there may be no neutral frame with which to express the information.<sup>9</sup> As the goal of this paper is to examine whether framing interferes with consent, I’ll take for granted my position that debiasing is not likely to work.<sup>10</sup> Additionally, if it turns out that framing effects do not interfere with consent, and if doctors’ main obligations are to ensure both the wellbeing and consent of their patients, then doctors are perhaps obligated to use framing effects to achieve the wellbeing of their patients. Overall, I argue that not only are framing effects unavoidable, there is no consent-based reason against using them. This suggests that, even if doctors could debias patients against framing, respect for informed consent does not require them to do so.

### **III. Why Framing Effects Do Not Undermine Consent**

Here, I argue that framing effects in general do not undermine consent. Then, in the next section, I argue that there is no compelling reason to think that the medical setting is different in a way that would make framing effects undermine consent. When I say “in general” I mean to say that this point is not limited to the medical context. I will canvas the plausible reasons why framing

---

<sup>8</sup> One might object that while framing effects engage system 1, ordering effects plus discussion of framing involve system 2, which is less prone to manipulation. Yet, there is evidence that ordering effects do change people’s decisions. To avoid this, would doctors need to disclose to patients about ordering effects? But would this re-introduce the problem at another level? Perhaps framing could be removed by saying both frames and then telling the patient about ordering effects. But, if I’m correct that framing doesn’t undermine consent and that doctors can use it to better fulfill their duty of beneficence, this would be irrelevant.

<sup>9</sup> Hanna 2011: 523. Chwang (2015: 8) also rejects this.

<sup>10</sup> See Chwang for a defense of debiasing.

effects might undermine consent and argue that they fail.<sup>11</sup>

### *III. 1. Framing effects undermine consent, because they deceive the person framed*

One might think that framing effects violate the requirements of informed consent, because they deceive the patient, thus violating the informed condition. However, by definition, framing effects cannot be deceptive. They are logically equivalent ways of stating the same information, just with a different emphasis.<sup>12</sup> Given this, there is no information contained in one frame that is not already contained in the other. As such, no information is withheld about the content of the treatment, and thus, no deception has occurred.<sup>13</sup>

### *III. 2. Framing effects undermine consent, because they withhold relevant background information from the person framed*

The person being framed is not being deceived about the content of what is being proposed to her. But, she is likely ignorant about whether framing effects are being used on her. Doctors who use framing effects on their patients may be failing at some obligation to disclose meta-information to their patients and are thus engaged in higher-order deception.

In non-medical contexts, it doesn't seem that the non-disclosure of meta-information about framing effects threatens consent. Consider this case:

**Furniture Sale:** S is a furniture salesperson and is trying to sell J a couch. S says (truthfully), “90% of these couches last over 5 years.” S uses this description (as opposed to the negative frame, “only 10% of these couches don't last over 5 years”), because he knows J is more likely to buy the couch with the positive frame. Without knowing that his behavior has been nudged, J agrees to buy the couch.

---

<sup>11</sup> Some of the authors I discuss are mainly concerned with framing in medicine and think that framing is unproblematic outside medicine. But, if framing effects undermine consent in medicine, then they should do so in non-medical situations. Hanna agrees with this: “The arguments I offer...apply to any case in which informed consent is generally thought necessary to waive rights” (Hanna 518).

<sup>12</sup> There is debate as to whether a positive and negative frame have identical semantic content. For a summary see Fisher (2022).

<sup>13</sup> Engelen endorses this view: “framing techniques risk making them [patients] less informed” (2019: 56).

It seems to me that S is under no obligation to disclose the second-order information about his use of framing, and it also seems clear that this nudge does not threaten J's consent to buy the couch. So long as people are not coerced, deceived and are otherwise competent, their wishes, framed or not, are morally transformative in commercial contexts. If we were to deny this, we would have to deny the consent in many commercial transactions, and we would have a very revisionary theory of consent. For example, transactions like buying a meal at a restaurant with framing effects on its display board would not be consensual.

I do not mean to endorse that the conditional "if it's permissible in business, then it's permissible in medicine." There are many things that are permissible in business but not in medicine. A car salesperson doesn't seem to act impermissibly if they use nudges to get me to add a bunch of expensive, unnecessary features to my car. But, were a doctor to convince me using nudges to get unnecessary procedures or a needlessly more expensive non-generic medicine, this would be wrong. Despite this, we can still infer from our consent practices in business to our consent practices in medicine. We just need to factor in the different role obligations that come with medicine. Doctors have a duty of beneficence that businesspeople (rarely) have. Doctors are obligated not just to not harm me but also to benefit me. Although we can reasonably disagree about the obligations of businesspeople to their clients, we can all agree that they do not have a duty to ensure that their customer gets the best deal, best product, etc. Businesspeople may be allowed to nudge us in whatever way they want, consistent with our consent, while doctors cannot nudge us in a way that conflicts with their duties of beneficence. Despite this difference in role obligations, we are still able to infer from consent practices in business to consent practices in medicine. The fact that doctors have a duty of beneficence to their patients that car salespeople don't have toward their customers isn't consent-relevant.

Whether a doctor has fulfilled their duty of beneficence is simply a different question. If A sells B a car, and we agree that it is consensual, we can infer that an analogous transaction in medicine would be consensual. A doctor benefiting a patient, while obligatory, is not consent-relevant. So long as the people involved are competent, the transaction is still consensual. It may be that, if the doctor fails her duty of beneficence, this interaction is wrong, but it would not be wrong for consent-based reasons.

### *III. 3. Framing effects undermine consent, because they manipulate the person framed*

One might argue that framing effects are manipulative. Without getting into the precise details of an account of manipulation, let's view manipulation roughly as an influence on an agent's decision that is not coercive, deceptive, nor persuasive but which often *seems* pro tanto wrong.<sup>14</sup> It seems plausible that framing effects are manipulative. They do not deceive or coerce the agent, but they seem to bypass her capacity for being persuaded. Some have argued that framing effects undermine consent due to being manipulative. There are two questions to hold separate: (1) are framing effects manipulative? And (2) if framing effects are manipulative, are they wrongfully manipulative in a way that would undermine consent?

Assume for the moment that framing effects are manipulative. While I feel the force of the claim that there is something pro-tanto wrong about manipulation, if (as some views hold) manipulation is not pro tanto wrong, then the fact the framing effects are manipulative would not tell against them. Instead, only framing effects that push the agent into a bad or harmful choice would be wrong.

---

<sup>14</sup> This usage of manipulation differs from the freewill literature. There, "the term 'manipulation' typically refers to radical programming or reprogramming of all or most of an agent's beliefs, desires, and other mental states. Such global manipulation...is also typically imagined as happening via decidedly extra-ordinary methods, such as supernatural intervention, direct neurological engineering, or radical programs of indoctrination and psychological conditioning" (Noggle 2020).

But, that is controversial. Assume for a moment that manipulation is always at least *pro tanto* wrong. I argue that, even if this is true, the unavoidability of framing effects renders their wrongness moot. Essentially, if framing effects are unavoidable in an interaction, then so is manipulation, meaning that it is not an option for one agent to not be manipulating the other in this interaction. If a doctor must manipulate a patient one way or the other, it seems that the wrong of manipulation is moot because it is unavoidable and that the doctor should manipulate the patient in a way that is in the patient's interest. To put it differently, if the doctor cannot help but nudge the patient, and the options are A (not in the patient's best interests) or B (in the patient's best interests), it seems obligatory for the doctor to frame the patient toward B. Surely, respect for consent cannot require us to push a patient toward something that is bad for them. This argument may strike many as too quick, but I don't think so. All it claims is that if a wrong-making feature is equally present in two actions, then this wrong-making feature should be a wash between the two of them. In other words, if two actions are equally wrong on the same metric, it doesn't make any sense to use that metric to decide between those actions. This is importantly different than the point sometimes made that because nudging is inevitable, it is therefore not manipulative to begin with. I am not saying that the inevitability removes the manipulation, only that it makes its wrongness moot. One might object with this case that I borrow (and slightly modify) from Noggle:

Suppose that Jones is traveling to a job interview on a subway car so crowded that it is inevitable that he will bump up against his fellow passengers. Suppose that he capitalizes on this fact to deliberately bump his rival job candidate (who is on the same subway car) out the door just as it closes, thus ensuring that he will be late for his interview. Clearly the fact that some bumping on Jones's part was inevitable does not excuse Jones's intentional bumping of his rival (2020).

But, if Jones must bump into someone, he should do so in a way that either does no harm or would benefit someone. So, all this proves is that we should not nudge into harm, not that we should not nudge.

Of course, doctors may be mistaken about a patient's best interests. I am assuming good-faith agents who take the patient's desires into account. In that manner, my view is a defense of framing a patient toward their good, defined as the medical option that best achieves what the patient wants. What the patient wants can also be shaped in the clinical setting via the process of shared decision making between doctors and patients.

Thus far, I've argued that if manipulation is always pro tanto wrong and if framing effects are manipulative, the fact that they are unavoidable renders the problem moot. But, suppose that framing effects were avoidable (perhaps via debiasing). Here, I canvas the dominant accounts of manipulation and show that either (1) framing effects are not manipulative on the given account or that (2) if framing effects are manipulative, they are not wrongly manipulative in a way that would threaten consent.

There are three main views about manipulation. Of necessity, this characterization will be broad. On the first view, "manipulation is...said to 'bypass,' 'undermine,' or 'subvert,' the target's rational deliberation" (Noggle 2020). For this view, framing effects are manipulative, because they introduce "non-rational influences into the deliberative process" (Noggle 2020).<sup>15</sup> But, this account of manipulation presents a standard where so many ordinary actions count as manipulative that it is hard to see why manipulation is pro tanto wrong. For example, Noggle points out that, on this view, appeals to the Golden Rule (as

---

<sup>15</sup> This view is also endorsed by Raz 1988: 377.

they draw on emotions) count as manipulative, insofar as they introduce non-rational features into the deliberative process. The general point is that this view counts things as manipulative which, intuitively, are not wrong. This suggests that, as an account of manipulation, this view permits of good and bad uses of manipulation. Presumably, nudging a patient toward something that is in her interests is a good use of manipulation. So, although framing effects would be manipulative, it would be in a sense that doesn't make them wrong.<sup>16</sup>

Manipulation is sometimes defined “as a form of trickery” (Noggle 2020).<sup>17</sup> More specifically, this view holds that an influence is manipulative if it induces a false belief in the agent. But, so long as the doctor is using framing effects to push a patient toward what is in her good (by the patient's own lights), then these framing effects would not count as manipulation. For example, if the doctor knows that the patient wants to live, knows that the procedure will achieve that goal, and also knows that the mortality rate frame would discourage the patient from getting the procedure, using the success rate frame would not count as manipulation at all, because the doctor is inducing a belief that the patient believes to be true and which objectively is true (i.e. “getting the surgery is right for me”).

On the final view, manipulation is “a kind of pressure to do as the influencer wishes” (Noggle 2020).<sup>18</sup> It seems that framing effects are manipulative on this account. But, this account counts too much as manipulation such that manipulation isn't always pro tanto wrong. Nudging (*not coercing*) someone into a choice that is good for them seems intuitively permissible,

---

<sup>16</sup> This view has counter-examples. See Gorin's (2014) case of manipulating someone into believing God doesn't exist by using rational arguments.

<sup>17</sup> This view is defended by Scanlon (1998) and in modified form by Cohen (2018).

<sup>18</sup> Faden, Beuchamp, King (1986), Conly (2004), and Buss (2005) defend a view like this. See Buss (2005: 210) and Conly (2004: 108-109) for support of this reading.

especially if not applying this nudge would predictably cause the person to endure an outcome that they don't want. So long as framing effects do that, then they would not count as wrongful manipulation.

In this section, I've argued that (1) if manipulation is not necessarily *pro tanto* wrong, then we shouldn't be concerned about whether framing effects are manipulative, (2) if framing effects are manipulative, then because framing effects are unavoidable, their wrongness is moot, and that (3) if framing effects are avoidable, then they are still not wrongfully manipulative, because they are either not manipulative or not wrongfully manipulative.<sup>19</sup>

One might argue that (2) is not promising, because whether the manipulation involved in framing is moot/inevitable, this does not bear on whether that manipulation undermines autonomy. To address this point, I will next argue that framing effects are not autonomy undermining.

#### *III.4. Framing effects undermine consent, because they violate the autonomy of the person framed*

One might argue that framing effects undermine autonomy and thus undermine consent. To address this, we must briefly digress into the literature on autonomy. The literature on autonomy is complex and difficult to parse. As I see it, there are two main categories of views: *substantive* and *procedural* views.<sup>20</sup> The former says that acting autonomously consists in having the proper relation between an individual and her desires. For example, views in this category have said that an agent is autonomous if she has coherence between her actions and her desires, if her desires and actions are authentic to her, if she identifies/resonates with her desires, if her first and

---

<sup>19</sup> For further argument that manipulation does not undermine consent, see Mandava and Millum (2013).

<sup>20</sup> My set-up is similar to Terlazzo (2016) and Stoljar (2014). Although this is not a perfect list, it captures enough. Relational and externalist views of autonomy can be subsumed into this dichotomy.

second-order desires cohere with each other, etc.<sup>21</sup> The substantive view of autonomy says that the content of one's preference determines whether it is autonomous. Thus, to determine if a choice is autonomous, we only need to know the content of the agent's desires, nothing else. The proceduralist view of autonomy (also called the historical view) says that acting autonomously consists in doing what one wants provided there has been no coercive intervention in the causal history of one's decision-making process.<sup>22</sup> I argue that framing effects (even if they are manipulative) are not a threat on either view of autonomy.

Start with the substantive view of autonomy. On this view, all that matters is the content of one's desire, not the causal process of arriving at that desire. Thus, as long as the framing effect nudges the agent toward the option that better fits her desires, then there is nothing autonomy undermining about framing. If a doctor were to nudge a patient away from the option that better fits with her desire, this would undermine her autonomy. But, it is not framing effects *qua* framing effects that would undermine the agent's autonomy in this case. Instead, it is *how* those framing effects are used by the doctor that determines if they undermine autonomy. The doctor could avoid this problem by attempting to discern what the patient wants and then framing her toward that option. So, yes, if we endorse the substantive view of autonomy, and if a doctor is framing a patient away from what that patient thinks is her good, then this would undermine her consent. But, this conclusion is compatible with framing effects actually being used to enhance a patient's autonomy.

To add briefly, the preceding conclusion is true even if framing effects are manipulative. On the substantive view of autonomy, being manipulated into the option that better fits with one's desires is autonomy enhancing.

---

<sup>21</sup> Frankfurt 1987), Wall (1998), and Raz (1988) defend this view.

<sup>22</sup> Christman (1991) is the most well-known defender of this view.

On the proceduralist account, what invalidates autonomy is a pernicious influence in the causal history of one's desires. The obvious candidate for this pernicious influence is coercion, but there are others. Framing effects are not coercive, but they may be a manipulative influence on one's desires. Several arguments suggest that framing effects do not undermine autonomy (and thus not consent) due to being manipulative. First, recall that on one of three main definitions of manipulation, framing effects are not manipulative (and they are not wrongfully manipulative on the remaining accounts). Second, while I feel the force of *global* manipulation (i.e. cases from the free will literature involving controlling background agents who manipulate one's whole environment) undermining autonomy, it seems too strong to say that *local* manipulation (i.e. interpersonal, conversational manipulation) can do so.<sup>23</sup> On the two views of manipulation according to which framing effects are manipulative (manipulation as pressure and manipulation as bypassing rational deliberation) it doesn't seem like manipulation must undermine consent. People can pressure me or provide non-rational influences into my decision process without undermining my consent. To suggest the contrary is highly revisionary. For example, if someone consents to sex due to the enticing cologne of a new partner or consents to a job offer when an erroneous feature of it is emphasized, consent is not undermined. So, it seems that a framing effect is not a strong enough influence in the causal history of an agent's desire to undermine her autonomy. To put this back into the context of medical framing effects, if an evil

---

<sup>23</sup> Christman thinks that manipulation can undermine autonomy. As I read Christman, he is discussing a global form of manipulation, not interpersonal-conversational manipulation. For example, in his discussion of manipulation undermining autonomy (1991: 19), he uses a case of a cult member who is indoctrinated, suggesting global manipulation in mind. In claiming that manipulation at the local level does not undermine autonomy, I follow Buss, who says that "I will argue...that when manipulation and deceit are morally criticizable, this is not because they undermine the autonomy of the one who is manipulated and deceived or even because this is what the manipulator/deceiver wills to do" (2005: 207).

doctor were to frame a patient against her best interests, this framing effect would not interfere with her procedural autonomy, although it would be wrong for other reasons.

### *III. 5. A rejection of Hanna's argument*

Here, I argue against Jason Hanna's view as to why framing effects undermine consent. Hanna argues that framing effects undermine consent, because they make consent or dissent depend on irrelevant factors. In response to this problem of triviality, Hanna says that the defender of consent can appeal to either the patient's hypothetical or actual consent as a way out, but neither option works by his estimation. I agree with Hanna that appealing to hypothetical consent is not a promising route, so I won't discuss it. Hanna thinks that the path involved in choosing actual consent is fraught. I disagree with this. Here is Hanna on actual consent:

A defender...may be tempted to offer the following resolution of this case: since Patient actually consents to surgery, it is permissible to proceed. Conversely, on this view, if Patient had actually dissented because she had been given the mortality rate description, then it would have been impermissible to proceed with surgery. In short, the proposal holds that treatment of Patient is morally governed by her actual choice (2011: 520).

In other words, we can appeal to what the patient actually consents to as indicative of her valid consent, even if we know that she would choose differently if she were told a different frame. But, this falls prey to the triviality problem. As Hanna says, "it is difficult to believe that the permissibility of an act could depend on something as trivial as how the question is asked" (2011: 520).

And, for Hanna, the triviality of actual consent under framing is even more problematic. On his view, we should want the patient to agree to the treatment under any morally relevant description of the treatment. He then argues that both the negative and positive frames (in our original surgery cases) are morally relevant descriptions for the patient. So, both the positive and negative frame are morally relevant for the patient, but the patient consents under one frame and

dissents under the other. This leads to the paradox that the patient has both waived and not waived her right to the treatment on different morally relevant descriptions of the same action.

As Hanna says:

It thus appears that Patient's actual choice does not determine the permissibility of proceeding...There are strong grounds for claiming, paradoxically, that Patient has waived her right (since she has consented under a morally relevant act description) and that she retains her right (since she is disposed to dissent under a morally relevant act description) (2011: 522).

My response is that Hanna takes too high a view of consent.<sup>24</sup> If consent is intrinsically important and is not just a useful guide to the agent's wellbeing, then surely it should not be so easily swayed? I disagree. To my mind, one can believe (1) that consent fulfills a robust gate-keeping function and that (2) people often give their consent on the basis of trivial factors. Hanna seems to suggest that the reality of (2) should make us doubt the truth of (1). This doesn't follow. Why think that consent has to be determined by substantive things? We consent to sex based on how someone smells, based on some odd remark someone made, etc. We are more likely to consent to future plans with friends if we are feeling happy than if we are feeling sad. A moment's reflection will reveal how often we consent based on extremely trivial features of a situation. I contend that this is compatible with the value and importance of informed consent. Hanna's view seems to rely on this principle:

**Incompatibility of Consent and Arbitrary Influence:** if an agent's consent or dissent counterfactually depends on an arbitrary influence (one that does not provide more information), then her consent is invalid.

This principle seems false given the numerous counter-examples to it. The effect of arbitrary influences is compatible with seeing autonomy and consent as importantly connected. On either

---

<sup>24</sup> There is an irony in paternalism (as well as other forms of skepticism) in that in the paternalist's argument for why consent doesn't matter very much, they must claim that consent is an extremely important and momentous phenomenon.

the substantive or procedural view, there is nothing necessarily consent undermining about an arbitrary influence on someone's decision. It doesn't coerce them or necessarily make them decide inauthentically to their desires, etc.

All of this sets up an explanation both for why Hanna is wrong and for why framing effects do not undermine consent. Because the Incompatibility of Consent and Arbitrary Influence thesis is false, framing effects do not threaten consent in virtue of allowing consent to depend on arbitrary influences. And, the fact that an agent would both consent and dissent to an action depending on how it is described (which is an arbitrary influence) does not undermine consent either. In the end, whether your consent depends on some arbitrary influence does not make it any less important. It is still what you decide, and is thus a reflection of your autonomous choice, making it worthy of respect. How odd would it be if a waiter said to a customer, "I'm not going to bring you the salad special, because had I mentioned the steak special instead, you would have picked it instead." Even though the customer's choice depended on an arbitrary feature of the situation (which special the waiter mentioned), that doesn't negate that she did make a choice that should be respected.

Perhaps Hanna could reply that the principle he relies on is more like the following:

**Incompatibility of Consent and Arbitrary Influence\*:** if an agent's consent or dissent counterfactually depends on an arbitrary influence (one that *ought* not be relevant to her decision), then her consent is invalid.

This thesis, I believe, would get Hanna closer to his desired conclusion. If it turns out that framing effects counterfactually influence one's decision to consent, and in a way that should not be relevant to her decision, then it seems more plausible to say that the value of autonomy is not being expressed by this consent. A trivial difference in wording that does not provide any different information ought not be relevant to the patient's decision. While that seems plausible,

the problem enters with this revised thesis. The revised thesis comes very close to saying that if an agent will behave in a practically irrational manner, the agent's consent is not valid and that the agent may be treated without regard to consent. This is the key point. As an anti-paternalist, I simply deny this. People can be inconsistent and practically irrational, and this does not license interference in their affairs. To suggest otherwise is to assert the truth of a particular kind of paternalism. It seems that paternalists and anti-paternalists are simply split on the question of whether arbitrary influences that are counterfactually relevant to a decision undermine consent.

Another way to think of this is that, for anti-paternalists, competent adults have the right to not be interfered with merely because they are being irrational. This is clearly a right that anti-paternalists believe in. Both the Incompatibility of Consent and Arbitrary Influence thesis and the revised version entail that competent adults do not have this right. To claim that is, in a way, to beg the question in favor of paternalism.

#### **IV. Medicine Is Not Special**

I've argued that there is no good reason to think that framing effects undermine consent in general. Here, I address whether there is something unique to the medical context that would make framing effects impermissible or consent undermining in a way that would not apply to the general question of consent. I argue that there is no such reason.<sup>25</sup>

The most obvious difference between the general context of consent and the medical context is the asymmetry of information between doctors and patients. Doctors know much more about medicine than patients do.<sup>26</sup> Does this informational asymmetry show that framing effects in the clinical setting are impermissible? I contend that it does not. Informational asymmetries

---

<sup>25</sup> Some argue that framing effects do not undermine consent in general but do undermine it in medicine (see Chwang 2015: 3).

<sup>26</sup> Chwang (2015: 3-4) addresses this point as well.

abound in contexts outside of medicine. Car salespeople, loan officers, realtors, etc. often know vastly more than their clients about their products. But, this doesn't seem to undermine or call into question the validity of these practices when framing effects are involved.

Perhaps the position could be revised to say that medicine is not special just because of the informational asymmetries between doctors and patients; it is special because of these asymmetries and the fact that doctors have a special duty to inform patients that other professions do not have. This seems plausible; a car salesperson doesn't have an obligation to disclose everything about the car to me, but the doctor has an obligation to disclose anything I want to know about the procedure to me. However, as I've argued earlier, the doctor's obligation to inform the patient is not violated by using a framing effect, because there is no information that the doctor is withholding from the patient on either frame. One might think that doctors are obligated to disclose meta-information about the existence of framing effects to patients. Earlier, I suggested that this duty to disclose meta-information was not present in business transactions that involve framing effects. But, perhaps it is obligatory for doctors to disclose this information. I think there is no consent-based reason to disclose this information, because this information itself doesn't eliminate the effect of the framing effect. Whether a patient knows that framing effects exist doesn't alter the effect of framing. It seems odd to suggest that doctors are obligated to disclose information that would not impact the decision of the patient on the grounds that it is relevant to the patient's consent. If it's no more relevant than the room number of the operating room, then there is no reason that respect for consent entitles the patient to know it.

Perhaps the unique feature of the medical domain is that it often involves high levels of risk, and the stakes are generally high. This may explain why framing effects are problematic in medicine and not in other contexts. But, being high-stakes is not unique to medicine, nor is it the

case for every medical situation. For example, buying a house is very high stakes, and realtors very likely use framing effects when discussing loan rates. This does not seem to undermine consent in those situations. Additionally, many medical situations are very low-stakes but may still involve framing effects. Something being high stakes is neither unique to medicine nor a reason to think that it would prohibit framing effects. This suggests that if framing effects and high stakes invalidate consent, they must do so equally in medicine and in other contexts.

Another difference could be that doctors have a special obligation not to manipulate patients that agents in general do not have. I feel the force of this, but I think that it still wouldn't prohibit the use of framing effects in medicine. If framing effects are inevitable, which I think they are, then doctors cannot avoid using them. Given this, doctors seem to be caught between their obligations not to manipulate patients and also their obligations of beneficence toward their patients. Regardless of which frame they use, they are exerting a manipulative influence on their patients. So, either frame is manipulative. However, one frame better fulfills the duty of beneficence than the other.<sup>27</sup> If the surgery is clearly good for the patient, and the doctor has no way to not manipulate the patient in one way or the other, it seems clear that the doctor should nudge the patient toward the beneficent decision. Thus, even if doctors have a special obligation not to manipulate patients (which it seems that they do), this will not ground a prohibition on framing effects. And, all of the previous arguments about whether framing effects are wrongful manipulation or even manipulative at all would apply here.

This leads me to another difference: doctors have a duty of beneficence to patients which is absent in market transactions. Unlike a car salesperson, my doctor is duty-bound to

---

<sup>27</sup> Some have argued that positive frames can reduce nocebo effects from medications (Barnes et al. 2019). A nocebo effect occurs when a patient hears about a negative side effect and is thus more likely to experience it (Cohen 2014: 147).

aggressively pursue what is best for me. This difference between medicine and markets would not ground a prohibition on framing effects, because part of a doctor pursuing my best interests may involve framing. If the doctor must frame a choice, the duty of beneficence demands that they frame it in a way that is medically best for the patient.

One final difference is that there is a power imbalance between doctors and patients. Doctors are authority figures, and humans tend to defer to authority. The nudge of a doctor and the nudge of a realtor do not come from the same position of authority.<sup>28</sup> Once again, there are non-medical transactions where this is also the case (i.e. lawyers, mechanics, plumbers, etc.).

I conclude that there is not a compelling reason to think that the medical context is uniquely different when it comes to the relationship between framing effects and consent.

All of this is compatible with my claim earlier that I do not mean to endorse the conditional “if it’s permissible in business, then it’s permissible in medicine.” However, I do think that the nature of consent does not change between each domain. If something is consensual in business, then (*ceteris paribus*) it is consensual in medicine. But, not all else is equal. Doctors have role obligations (to inform and to be beneficent) in a way that is rarely true in market transactions. In this section, I’ve simply added to this by saying that there is no feature of medicine (which is not present in business) that would uniquely tell against framing effects.

## **V. Is There a Non-Consent-Based Defense of Disclosure?**

I have argued that there is no consent-based reason to prohibit framing effects in medicine or in general. So, what are we left with? Should doctors rampantly frame their patients without telling them that they are doing so? I think that the answer to this question is ‘yes,’ with important reservations. One might have this worry about using framing effects, even if they don’t conflict

---

<sup>28</sup> A similar point is made by Engelen and Ploug et al (2019).

with consent: imagine if the literature on nudging were to become widely read by the general public, and patients were to see defenses of secret nudging from doctors to patients. Were this information to become widespread, those people might begin to lose trust in medicine.<sup>29</sup> To prevent implications such as this, the objector argues that doctors should make patients aware of the existence of framing effects and their use in medicine. Doctors may still use framing effects, but they must tell patients that they are doing so.

While I'm sympathetic to this concern, the empirical evidence is scant and mixed. If it turns out that disclosing the use of nudges has no effect on trust, then there would be no reason for disclosure, at least no more reason than there is to disclose something like the room number of the operating room. A 2015 interview study (N=21) suggests that this may be true, finding that "consumers are generally appreciative of nudging both as a general concept and when targeting health behaviors" (Junghans et al., 2015: 11). A possible explanation for this trust in nudging comes from the fact that the people being nudged had positive attitudes toward the nudging agent or institution and a trust that it/they had the patient's best interests in mind (Junghans et al., 2015: 8). As the authors find, "interviewees generally expressed that if the intention behind the nudge was good, as most agreed in the case of health behavior and healthy eating, they would not be particularly concerned with the actors who design or implement the nudges" (Junghans et al., 2015: 9). If patients trust their doctors and think that their doctors have good intentions, this evidence suggests that they are favorably disposed to being nudged, which in turn suggests that finding out they had been nudged would not undermine their trust in doctors.

---

<sup>29</sup> Wilkinson (2017) raises a similar question.

Now, to my reservations about not disclosing meta-information about the use of framing. From the same interview, some patients expressed a negative attitude to retroactively finding out they had been nudged:

Interviewees also expressed that they would not appreciate if they realized that they had been led to a decision that was out of their awareness. This did not necessarily mean that they did not want to be nudged, but if so, they did not want to detect the influence.

In short, there is some reason to believe that covert nudging is unstable, because patients would lose trust if they found out about it. Given the small sample, more research should be done before concluding one way or the other on this point.

As of now, I do not think there is sufficient reason to disclose this information on the grounds of maintaining trust, and there is no consent-based reason to disclose. Thus, doctors may use framing effects to nudge patients toward their good, with the caveat that framing should be disclosed if more evidence emerges that patients would lose trust in doctors were they to find out about covert nudges.

Although this paper is about a topic in bioethics, it has broader implications. If I am correct, it will turn out that (1) a very common form of nudging is compatible with consent and autonomy broadly (in and outside of medicine), (2) that the standard for consent in medicine and business transactions is the same (although there are different role obligations in each domain that may change the rules).

## References

- Barnes, Kirsten, et al. (2019) Can Positive Framing Reduce Nocebo Side Effects? Current Evidence and Recommendation for Future Research. *Frontiers in Pharmacology* 10: 167: 1-9.
- Bermúdez, José Luis (2022) Rational framing effects: A multidisciplinary case. *Behavioral and Brain Sciences*: 1-67.
- Blumenthal-Barby, J.S. (2021) *Good Ethics and Bad Choices: The Relevance of Behavioral Economics for Medical Ethics*. MIT Press.
- Buss, Sarah (2005) Valuing Autonomy and Respecting Persons: Manipulation, Seduction, and the Basis of Moral Constraints. *Ethics* 115: 195-235.
- Chwang, Eric (2015) Consent's Been Framed: When Framing Effects Invalidated Consent and How to Validate It Again. *Journal of Applied Philosophy* 33: 3: 1-16.
- Christman, John (1991) Autonomy and Personal History. *Canadian Journal of Philosophy* 21: 1: 1-24.
- Cohen, Shlomo (2013) Nudging and Informed Consent. *The American Journal of Bioethics* 13: 6: 3-11.
- (2018) The Nocebo Effect of Informed Consent. *Bioethics* 28: 3: 147-154.
- (2018) Manipulation and Deception. *Australasian Journal of Philosophy* 96: 3: 483-497.
- Conly, Sarah. (2004) Seduction, Rape, and Coercion. *Ethics* 115: 96-121.
- Engelen, Bart (2019) Ethical Criteria for Health-Promoting Nudges: A Case-by-Case Analysis. *The American Journal of Bioethics* 19: 5: 48-59.
- Faden, Ruth R., Tom L. Beauchamp, and Nancy M.P. King (1986) *A History and Theory of Informed Consent*. New York: Oxford University Press.

- Fisher, Sarah (2022) Meaning and framing: the semantic implications of psychological framing effects. *Inquiry*: 1-25.
- (2020) RATIONALISING FRAMING EFFECTS: AT LEAST ONE TASK FOR EMPIRICALLY INFORMED PHILOSOPHY. *CRÍTICA* 52: 156: 5-30.
- Frankfurt, Harry (1987) "Freedom of the Will and the Concept of a Person," In: *The Importance of What We Care About*, Cambridge: Cambridge University Press, pp.11–25.
- Gonzalez, Cleotilde, et al. (2005) The Framing Effect and Risky Decisions: Examining Cognitive Functions with fMRI. *Journal of Economic Psychology* 26: 1-20.
- Gorin, Moti (2014) Do Manipulators Always Threaten Rationality? *American Philosophical Quarterly*, 51: 1: 51-61.
- Hanna, Jason (2011) Consent and the Problem of Framing Effects. *Ethical Theory and Moral Practice* 14: 517-531.
- Junghans, Astrid, et al. (2015) Under consumers' scrutiny - an investigation into consumers' attitudes and concerns about nudging in the realm of health behavior. *BMC Public Health* 15: 336: 1-13.
- Levy, Neil (2020) "Nudge, Nudge, Wink, Wink: Nudging is Giving Reasons." *Ergo* 6: 10: 281-302.
- Mandava, Amuyla, and Joseph Millum (2013) Manipulation in the Enrollment of Research Participants. *Hastings Center Report* 43: 2: 38-47.
- McNeil B, Pauker S, Sox H, and Tversky A (1982) On the elicitation of preferences for alternative therapies. *New England Journal of Medicine* 306: 1259–1262.
- Noggle, Robert (2020) "The Ethics of Manipulation." *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/ethics-manipulation/2020>.

- Ploug, T. et al. (2012) To nudge or not to nudge: cancer screening programmes and the limits of libertarian paternalism. *Journal of Epidemiology and Community Health* 66: 12: 1193-1196.
- Raz, Joseph (1988) *The Morality of Freedom*, Oxford: Clarendon.
- Scanlon, T.M (1998) *What We Owe to Each Other*. Cambridge, MA: Harvard University/Belknap Press.
- Simkulet, William (2017) Nudging, informed consent, and bullshit. *Journal of Medical Ethics* 44: 536-542.
- Stoljar, Natalie (2014) Autonomy and Adaptive Preference Formation. In: *Autonomy, Oppression, and Gender*. Edited by Veltman and Piper. Oxford. Pp. 227-254.
- Terlazzo, Rosa (2016) Conceptualizing Adaptive Preferences Respectfully: An Indirectly Substantive Account. *The Journal of Political Philosophy* 24: 2: 206-226.
- Thaler, Richard and Cass Sunstein. 2009. *Nudge*. Penguin.
- Tversky, Amos Tversky and Daniel Kahneman (1981) The framing of decisions and the psychology of choice.” *Science* 211: 4481: 453–458.
- Wall, Steven (1998) *Liberalism, Perfectionism and Restraint*. New York: Cambridge University Press.
- Wilkinson, T.M. (2017) Counter-Manipulation and Health Promotion. *Public Health Ethics* 10: 3: 257-266.