



# Fifty Shades of Manipulation

## Citation

Cass R. Sunstein, Fifty Shades of Manipulation, J. Behavioral Marketing (forthcoming 2015).

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:16149947>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

## **Fifty Shades of Manipulation**

**Cass R. Sunstein\***

### **Abstract**

*A statement or action can be said to be manipulative if it does not sufficiently engage or appeal to people's capacity for reflective and deliberative choice. One problem with manipulation, thus understood, is that it fails to respect people's autonomy and is an affront to their dignity. Another problem is that if they are products of manipulation, people's choices might fail to promote their own welfare, and might instead promote the welfare of the manipulator. To that extent, the central objection to manipulation is rooted in a version of Mill's Harm Principle: People know what is in their best interests and should have a (manipulation-free) opportunity to make that decision. On welfarist grounds, the norm against manipulation can be seen as a kind of heuristic, one that generally works well, but that can also lead to serious errors, at least when the manipulator is both informed and genuinely interested in the welfare of the chooser.*

*For the legal system, a pervasive puzzle is why manipulation is rarely policed. The simplest answer is that manipulation has so many shades, and in a social order that values free markets and is committed to freedom of expression, it is exceptionally difficult to regulate manipulation as such. But as the manipulator's motives become more self-interested or venal, and as efforts to bypass people's deliberative capacities becomes more successful, the ethical objections to manipulation become very forceful, and the argument for a legal response is fortified. The analysis of manipulation bears on emerging first amendment issues raised by compelled speech, especially in the context of graphic health warnings. Importantly, it can also help orient the regulation of financial products, where manipulation of consumer choices is an evident but rarely explicit concern.*

### **I. Introduction**

It ranks among the most powerful scenes in the history of television. Don Draper is charged with producing an advertising campaign for Kodak, which has just invented a new slide projector, with continuous viewing. It operates like a wheel. Using the device

---

\* Robert Walmsley University Professor, Harvard University. I am grateful to Oren Bar-Gill, Anne Barnhill, Tyler Cowen, Elizabeth Emens, Nir Eyal, George Loewenstein, Martha Minow, Martha Nussbaum, Eric Posner, Lucia Reisch, Mark Tushnet, and Adrian Vermeule for valuable comments on a previous draft. Lisa Marrone provided excellent research assistance.

to display scenes from a once-happy family (as it happens, his own, which is now broken), Draper tells his potential clients<sup>1</sup>:

*In Greek, "nostalgia" literally means, "the pain from an old wound." It's a twinge in your heart, far more powerful than memory alone. This device isn't a spaceship. It's a time machine. It goes backwards, forwards. It takes us to a place where we ache to go again. It's not called the Wheel. It's called a Carousel. It lets us travel the way a child travels. Around and around, and back home again . . . to a place where we know we are loved.*<sup>2</sup>

The Kodak clients are sold; they cancel their meetings with other companies.

Now consider the following cases:

1. A parent tries to convince an adult child to visit him in a remote town in Nebraska, saying, "After all, I'm your father, and I raised you for all those years, and it wasn't always a lot of fun for me -- and who knows whether I'm going to live a lot longer?"
2. An automobile company advertises its new vehicle by showing a sleek, attractive couple exiting from it before going to a glamorous party.
3. In an effort to discourage people from smoking, a government requires cigarette packages to contain graphic, frightening health warnings, depicting people with life-threatening illnesses.<sup>3</sup>
4. In a campaign advertisement, a political candidate displays highly unflattering photographs of his opponent, set against the background of frightening music. An announcer reads quotations that, while accurate and not misleading, are taken out of context to make the opponent look at once ridiculous and scary.

---

<sup>1</sup> See <http://www.imdb.com/title/tt1105057/quotes>

<sup>2</sup> Revealingly, nostalgia actually means "longing for a return home," rather than "pain from an old wound."

<sup>3</sup> On the FDA's effort to require graphic warnings on packages, see *RJ Reynolds v. FDA*, 696 F.3d 1205 (DC Cir 2012), available at [http://www.cadc.uscourts.gov/internet/opinions.nsf/4C0311C78EB11C5785257A64004EBFB5/\\$file/11-5332-1391191.pdf](http://www.cadc.uscourts.gov/internet/opinions.nsf/4C0311C78EB11C5785257A64004EBFB5/$file/11-5332-1391191.pdf), affirmed on other grounds, *RJ Reynolds v. FDA*, F3d (DC Cir. 2012), available at [http://www.cadc.uscourts.gov/internet/opinions.nsf/4C0311C78EB11C5785257A64004EBFB5/\\$file/11-5332-1391191.pdf](http://www.cadc.uscourts.gov/internet/opinions.nsf/4C0311C78EB11C5785257A64004EBFB5/$file/11-5332-1391191.pdf).

For the government's own graphic campaign, see <http://www.cdc.gov/tobacco/campaign/tips/resources/videos/> For evidence of success from graphic warnings, see [http://www.cdc.gov/media/releases/2011/p0526\\_cigarettewarnings.html](http://www.cdc.gov/media/releases/2011/p0526_cigarettewarnings.html)

5. In an effort to convince consumers to switch to its new, high-cost credit card, a company emphasizes its very low “teaser rate,” by which consumers can enjoy low-cost borrowing for a short period. In its advertisement, it depicts happy, elegant, energized people, displaying their card and their new purchases.
6. To reduce pollution (including greenhouse gas emissions), a city requires public utilities to offer clean energy sources as the default providers, subject to opt-out if customers want to save money.<sup>4</sup>

Both public and private law are pervasively concerned with the problem of coercion, arising from the literal use of force. The Due Process Clause is designed to impose procedural safeguards in the event of actual or threatened coercion on the part of government, and if private actors plan to resort to force, both criminal law and the law of tort will stand in their way. There are also legal constraints on lying and deception.<sup>5</sup> The first amendment protects commercial advertising, but it does not ban regulation of false or deceptive commercial speech.<sup>6</sup> The Federal Trade Commission is explicitly authorized to control “unfair and deceptive” trade practices.<sup>7</sup>

But what of manipulation, undertaken by either private or public institutions? Within law, there is no sustained analysis of the topic.<sup>8</sup> To be sure, there is a great deal of work on lies and deception,<sup>9</sup> and we can identify an overlap among lying, deceiving, and manipulating. We could even see manipulation as a master concept that includes lying and deceiving, or understand the three to be on some kind of continuum. Certainly this is so if our master principle is personal autonomy; if so, the three violate that principle,

---

<sup>4</sup> See Cass R. Sunstein and Lucia Reisch, *Automatically Green: Behavioral Economics and Environmental Protection*, 38 *Harv Env L Rev* 127 (2014). The example should not be taken to suggest that green energy is necessarily more expensive than other forms.

<sup>5</sup> Within Anglo-American law, deceit has long been tortious. See *Derry v. Peek*, LR 14 App Cas 337 (1889); John Hannigan, *Measure of Damages in Tort for Deceit*, 18 *B.U. L. Rev.* 681 (1938). An extensive body of law deals with the related issue of misrepresentation. See John Cartwright, *Misrepresentation, Mistake and Non-disclosure* (3d ed. 2012). On the ethical issues associated with lying, see Sissela Bok, *Lying* (2011).

<sup>6</sup> See *Virginia State Pharmacy Board v. Virginia Citizens Consumer Council*, 425 US 748 (1976).

<sup>7</sup> 15 USC 57a (a)(1)(B).

<sup>8</sup> The most valuable treatment involves the analogous problem of deception. Richard Craswell, *Interpreting Deceptive Advertising*, 65 *BU L Rev* 657 (1985); *Regulating Deceptive Advertising: The Role of Cost-Benefit Analysis*, 64 *Southern Cal. L. Rev.* 549 (1991). As we shall see, manipulation is a different concept, and it is much harder to define and police. Craswell’s superb discussions nonetheless bear on the question of regulating manipulation, and indeed helps show why regulation is so difficult.

<sup>9</sup> See Craswell, *Interpreting Deceptive Advertising*, *supra* note, for the seminal discussion.

though for somewhat different reasons. (I shall have something to say about the extent to which this is so.) But in ordinary usage, it is reasonable to think that the concept of manipulation is distinctive, certainly in the sense that it can occur (as in the mythical Kodak commercial) without lies or deception (at least in their standard forms). What does manipulation entail, and what is wrong with it? How, if at all, should the law respond to it – for example, in the context of consumer protection or in constraining government itself? One of my major goals here is to put the category of manipulation on the legal viewscreen.

It should be clear that an action does not count as manipulative merely because it is an effort to alter people's behavior. If you are a passenger in a car, and you warn the driver that he is about to get into a crash, you are not engaged in manipulation. The same is true if you remind someone that a bill is due. A calorie label and an energy efficiency label are not ordinarily counted as forms of manipulation. So long as a private or public institution is informing people, or "just providing the facts," it is hard to complain of manipulation.<sup>10</sup> There is also a large difference between persuading people and manipulating them. With (non-manipulative) persuasion, people are given facts and reasons, presented in a sufficiently fair and neutral way; manipulation is something different.

It is often thought that when people are being manipulated, they are treated as "puppets on a string."<sup>11</sup> Almost no one wants to be someone else's puppet (at least without consent<sup>12</sup>), and in some respects, it is especially bad to be a puppet of government. Many of the worst governments in history have attempted to turn their citizens into puppets.<sup>13</sup> If we keep the puppet metaphor in mind, the idea of "manipulation" can be applied to many kinds of behavior; but it is not entirely clear that it is a unitary concept, or that we can identify necessary and sufficient conditions.<sup>14</sup>

---

<sup>10</sup> A qualification is necessary. If a disclosure requirement focuses on one of many aspects of a situation, and fixes people's attention on that aspect, a charge of manipulation would not be unreasonable. Consider the controversy over the idea that sellers should have to disclose that food has genetically modified organisms (GMOs). See Charles Noussair et al., Do Consumers Really Refuse to Buy Genetically Modified Food?, 114 *Economic Journal* 102 (2004). For those who object to compulsory labeling about GMOs, there is a plausible claim that labels are a form of manipulation, activating public concern where there is no objective reason for that concern. Of course those in the private sector might engage in similar forms of manipulation, drawing people's attention to a feature of a product that, while real, appears far more important than it actually is.

<sup>11</sup> See T. M. Wilkinson, Nudging and Manipulation, 61 *Political Studies* 341, 242 (2013).

<sup>12</sup> See below.

<sup>13</sup> See Frank Westerman, *Engineers of the Soul: The Grandiose Propaganda of Stalin's Russia* (2012); Susan Bachrach, *State of Deception: The Power of Nazi Propaganda* (2009).

<sup>14</sup> For a number of instructive treatments, see *Manipulation* (Christian Coons and Michael Webster eds. 2014).

Manipulation takes multiple forms. It has at least fifty shades, and some people wonder if they are tightly identified with one another.<sup>15</sup>

The principal goal of this Article is to make progress in understanding what manipulation is and what is wrong with it. If we can make progress on those tasks, we should be better equipped to assess a wide range of problems in ethics, policy, and law. For example, plausible objections might be made to acts of government that can be counted as manipulative; such objections might not treat citizens with respect.<sup>16</sup> There are also first amendment questions. When the government compels speech, is there a difference between mandating a purely factual disclosure and mandating one that has arguably manipulative features? When is advertising manipulative, and if it is, what, if anything, should be done about it<sup>17</sup>? Are there circumstances in which manipulative speech, on the part of government, raises due process problems<sup>18</sup> or runs afoul of existing statutes<sup>19</sup>? How, if at all, should government respond to manipulative behavior by the

---

<sup>15</sup> A valuable cautionary note: “People can be manipulated when they go shopping, strike contracts, vote, study at school, visit their doctors, decide whether to have sex or take turns to do the housework. A full account would have to cope with the enormous variety of sites and methods of manipulation. Indeed, we do not have such an account.” Wilkinson, *supra* note, at 344.

<sup>16</sup> See Jeremy Waldron, *It’s All For Your Own Good*, *New York Review of Books* (2014), available at <http://www.nybooks.com/articles/archives/2014/oct/09/cass-sunstein-its-all-your-own-good/> Consider in particular this passage: “Deeper even than this is a prickly concern about dignity. What becomes of the self-respect we invest in our own willed actions, flawed and misguided though they often are, when so many of our choices are manipulated to promote what someone else sees (perhaps rightly) as our best interest?” In the particular context that concerns Waldron, I believe that the question is overheated, see Cass R. Sunstein, *Nudges and Choice Architecture: Ethical Considerations*, *Yale J Reg* (forthcoming 2015) (arguing that nudges are generally not manipulative, and should not be adopted if they are), but the question is legitimate in many other contexts and in the abstract. An instructive discussion is Pelle Goldberg Hansen and Andreas Maaloe Jespersen, *Nudge and the Manipulation of Choice*, 4 *EJRR* 3 (2013).

<sup>17</sup> For a provocative discussion, see Allen W. Wood, *Coercion, Manipulation, Exploitation*, in *Manipulation*, *supra* note, at 17.

<sup>18</sup> A possible “yes” answer is provided in *Wickard v. Filburn*, 317 US 111 (1942), though the Court ruled “no” on the particular facts, where the Secretary of Agriculture gave an arguably manipulative speech on behalf of a referendum: “There is no evidence that any voter put upon the Secretary’s words the interpretation that impressed the court below or was in any way misled. There is no showing that the speech influenced the outcome of the referendum.” *Id.* at 116.

<sup>19</sup> Labor law has an important pocket of doctrine that raises this question, though the fundamental problem is coercion (in the form of threats) rather than manipulation. See *NLRB v. Gissel Packing Co.*, 395 US 575 (1969).

private sector,<sup>20</sup> for example in the context of financial products (such as credit cards or mortgages)? As we shall see, an understanding of manipulation bears directly on the work of the Consumer Financial Protection Bureau (CFPB), and might help to orient some of its work,<sup>21</sup> which is at least implicitly concerned with the problem.

I suggest that an effort to influence people's choices counts as manipulative *to the extent that it does not sufficiently engage or appeal to their capacity for reflection and deliberation*. The word "sufficiently" leaves a degree of ambiguity and openness, and properly so. It is not possible to know whether manipulation is involved without asking about the sufficiency of people's capacity to deliberate on the question at hand.<sup>22</sup> We can imagine clear cases of manipulation (subliminal advertising<sup>23</sup>), cases that clearly fall outside of the category (for example, a warning about deer crossings in a remote area), and cases that can be taken as borderline (a vivid presentation about the advantages of a particular mortgage or a redesign of a website to attract customers to the most expensive products).<sup>24</sup> It is important to emphasize that countless choices are at least partly a product of variables that do not involve reflective deliberation – and choosers tend to be unaware of that fact.<sup>25</sup> The problem of manipulation arises when choosers justly complain that because of the actions of a manipulator, they have not, in a sense, had a fair chance to make a decision on their own.<sup>26</sup> Often the distinguishing mark of manipulation is a justified sense of ex post betrayal.

---

<sup>20</sup> A separate body of law deals with manipulative behavior in connection with swaps; the "manipulation" is a term of art in that setting. See <http://www.sec.gov/rules/proposed/2010/34-63236.pdf>

<sup>21</sup> See below; see also the discussion of potential regulation of payday loans in <http://www.wsj.com/articles/cfpb-sets-sights-on-payday-loans-1420410479>

<sup>22</sup> Compare the related discussion in Anne Barnhill, *What is Manipulation?* in *Manipulation: Theory and Practice* 50, 72 (Christian Coons and Michael Weber eds. 2014).

<sup>23</sup> See Augustus Bullock, *The Secret Sales Pitch: An Overview of Subliminal Advertising* (2004).

<sup>24</sup> Importantly, the word "sufficiently" applies to the degree of reflection and deliberation that are involved; it does not speak to the issue of justification. For example, would-be kidnappers might be manipulated (in the sense that their deliberative capacities are bypassed) by police officers who are trying to stop a kidnapping, and a terrorist might similarly be subject to (justified) manipulation.

<sup>25</sup> See Daniel Kahneman, *Thinking, Fast and Slow* (2011). For a related point from a different angle, see Edna Ullmann-Margalit, *Big Decisions: Opting, Converting, Drifting*, 81 *Royal Institution of Philosophy Supp.* 81 (2006) (contending that for small and large decisions, choices might not be based on reasons).

<sup>26</sup> There is, however, a set of cases that complicate the definition I offer here, and that suggest that it does not exhaust the category of manipulation. Suppose that people's judgments are properly and legitimately automatic and not a product of deliberation. (Immediate attractions to certain foods or persons are plausible examples.) We can imagine efforts to alter those automatic judgments through rational arguments that cannot be characterized as manipulative. But we can also imagine efforts to alter those

Of course there are degrees of manipulation, as some forms of influence attempt to bypass deliberation altogether (such as subliminal advertising), and other forms merely try to influence it by triggering certain forms of automatic processing (for example, through framing a problem so as to provoke the desired response<sup>27</sup>). Some forms of manipulation are modest and relatively benign. In the Kodak commercial, the goal is to connect the product with a set of evocative associations – childhood, a carousel, and a magical ability to recapture, and make permanent, a lost past. Manipulation often occurs through such associations, which are a pervasive feature of political campaigns<sup>28</sup>; but the concept is much broader than that. Manipulators often describe choices so as to make certain outcomes vivid and appealing (such as purchases of lottery tickets), or vivid and unappealing (such as failures to buy life insurance) -- even though a more neutral frame would present the whole problem in a less tendentious manner, leaving the chooser in a more objective position to weigh the relevant variables (and in that sense more free).

In such cases, a central problem with manipulation is that it can violate people's autonomy (by making them instruments of another's will) and offend their dignity (by failing to treat them with respect).<sup>29</sup> The manipulator is leading the chooser to make a choice without sufficiently assessing, on the chooser's own terms, its costs and its benefits. For this reason, the most strongly felt moral objections to manipulation are deontological in character. The objections reflect a sense that people are not being treated respectfully. Their own capacities – to assess, to weigh, to judge -- are not being given appropriate deference. For deontologists, a central question is whether choosers have given appropriate consent to the act of manipulation, or whether the manipulator has properly inferred consent under the circumstances.

From the welfarist point of view, the objection to manipulation is much less straightforward. Some people can benefit (a great deal) from being manipulated (consider a smoker who desperately wants to quit), and within limits, being manipulated can even

---

judgments that do not involve rational arguments at all. A friend, or an outsider, might attempt to use associations, or vivid pictures of some kind, to create a relevant change. The question is: Mightn't such cases involve manipulation, even if they do not involve judgments that ought to involve reflection and deliberation? That question raises the possibility that nondeliberative efforts to alter properly nondeliberative judgments might also be counted as manipulative. But discussion of this possibility would take me beyond my focus here. I am grateful to Anne Barnhill for raising this point, and see her discussion in Anne Barnhill, *What is Manipulation?* in *Manipulation: Theory and Practice* 50 (Christian Coons and Michael Weber eds. 2014).

<sup>27</sup> See *Perspectives on Framing* (Gideon Keren ed. 2010).

<sup>28</sup> The most famous, or infamous, example is Lyndon Johnson's terrifying "mushroom cloud" advertisement in his 1964 campaign against Barry Goldwater, ending, "The stakes are too high for you to stay home." See

<https://www.youtube.com/watch?v=uO0R4k1tVMs>

<sup>29</sup> The "undue influence" doctrine in contract raises related concerns. See Ray Madoff, *Unmasking Undue Influence*, 81 *Minn L Rev* 571 (1997).



be fun. In some forms, manipulation is a form of play, undertaken with a smile and a wink. (A speculation: Those who are intensely opposed to manipulation, in all its shades and forms, lack a sense of humor.) In other forms, it is not fun at all, even deadly serious (consider efforts to manipulate kidnappers or terrorists<sup>30</sup>). On welfarist grounds, there is no simple evaluation of manipulation, at least if we embrace the foregoing definition.

The foundation of the welfarist concern, I suggest, is the view, associated with Mill<sup>31</sup> and Hayek,<sup>32</sup> that the chooser, and not the manipulator, knows what is in his best interest. Of course Mill's principal concern, and Hayek's too, is with coercion, but the welfarist objection to manipulation stems from the same source: a belief that choosers know best. It follows that the anti-manipulation principle is strongly derivative of Mill's Harm Principle; it suggests that choosers ought to be able to make their own decisions, and that the role of others should be restricted to informing them or attempting to persuade them (without manipulation).<sup>33</sup>

If choosers know best, then the welfare-increasing approach is to avoid manipulation and to engage the chooser's deliberative capacities. But the manipulator refuses to do that. The skeptic wonders: Why not? A tempting answer is that the manipulator is promoting his own interests, and not those of the chooser. The use of manipulation, rather than (say) information or persuasion, creates a risk that the manipulator does not have the chooser's interests in mind. For that reason, manipulation undermines the welfare of the chooser. The welfarist analysis of manipulation closely parallels the welfarist analysis of fraud and deceit.<sup>34</sup> In a sense, the manipulator can even be seen as a kind of thief, taking something from the chooser without real consent. In some cases, that is indeed the right way to assess an act of manipulation; it helps to illuminate recent initiatives in the area of consumer financial protection.<sup>35</sup>

From the standpoint of the legal system, the problem is that as defined here, manipulation can plausibly be said to be pervasive. It can be found on television, on the Internet, in every political campaign, in countless markets, in friendships, and in family life. Even if we insist (as we should) that manipulation cannot occur without intentional manipulators,<sup>36</sup> the scope of the practice is very wide. It would be odd and perhaps

---

<sup>30</sup> For relevant discussion, see Cass R. Sunstein and Adrian Vermeule, *Conspiracy Theories: Causes and Cures*, 17 *J. Polit. Phil.* 202 (2009).

<sup>31</sup> See John Stuart Mill, *On Liberty* (1859).

<sup>32</sup> See Friedrich Hayek, *The Market and Other Orders* (2014).

<sup>33</sup> It is true, however, that people might consent to manipulation, just as they might consent to coercion. See Jon Elster, *Sour Grapes* (1983); John Beshears et al., *Self Control and Liquidity: How to Design A Commitment Contract* (unpublished manuscript 2015). For discussion of consent, see *infra*.

<sup>34</sup> See Pierre Picard, *Economic Analysis of Insurance Fraud*, in *Handbook of Insurance* 349 (2013).

<sup>35</sup> See *infra*.

<sup>36</sup> Nature can, in a sense, manipulate people; cold weather and snow, for example, can affect people without sufficiently triggering deliberation. But it seems useful to limit the

pointless to condemn practices that people encounter daily, and with which they live while mounting little or no objection.<sup>37</sup> Indeed, it would be fussy and stern – even a bit inhuman – to try to excise it. Because of the pervasiveness of manipulation, and because it often does little or no harm, the legal system usually does not attempt to prevent it.<sup>38</sup> At least in general, the costs of regulating manipulation would far exceed the benefits. But as we shall see, the proper evaluation of acts of manipulation depends a great deal on context, including the expectations associated with particular roles.<sup>39</sup> In some contexts, regulators do aim at manipulation, at least implicitly.<sup>40</sup>

Everyone knows that a car company wants to sell cars, and under existing conventions, it is acceptable to produce advertisements that do not exactly target people’s deliberative capacities (at least if falsehoods are not involved). Something similar can be said about political campaigns. To be sure, there remains a question whether deliberative capacities are “sufficiently” engaged; in such cases, they are hardly on hold. But it would be plausible to suggest that the sufficiency requirement is not met. If so, the ethical objection gains strength under two conditions: (1) when the manipulator’s goals are self-interested or venal and (2) when the act of manipulation is successful in subverting or bypassing the chooser’s deliberative capacities. When both conditions are met, there is good reason for an ethical taboo on manipulation, and perhaps legal constraints, certainly (but not only) when it comes from government. As we shall see, there is also reason for heightened concern, from the standpoint of the first amendment, when the government compels speech in order to manipulate those who encounter it; in such cases, the government should face an elevated burden of justification.

The remainder of this Essay is organized as follows. Part II explores conceptual and definitional issues, distinguishing manipulation from deceit, and linking the idea of

---

category to intentional efforts; in ordinary language, intentionality appears to be a defining characteristic of the concept of manipulation.

<sup>37</sup> I bracket the existence of longstanding and entrenched practices, such as discrimination on the basis of sex and disability, which are rightly subject to scrutiny and objection.

<sup>38</sup> For exceptions, see below.

<sup>39</sup> On the relationship between “nudges” and manipulation, see Wilkinson, *supra* note; Cass R. Sunstein, *Nudging and Choice Architecture: Ethical Considerations*, Yale J Reg (forthcoming 2015). In my view, nudges generally do not count as manipulative in the sense defined here. Consider, for example, information disclosure, reminders, and warnings, none of which is plausibly treated as manipulative (in the absence of special circumstances). Graphic warnings, loss aversion, default rules, and social norms are taken up below. It is true, however, that some nudges (not defended by me or to my knowledge anyone else) can cross the line; for general discussion, see *id.*

<sup>40</sup> An example involves a 2014 rule requiring integrated mortgage disclosures, available at [http://files.consumerfinance.gov/f/201311\\_cfpb\\_final-rule\\_integrated-mortgage-disclosures.pdf](http://files.consumerfinance.gov/f/201311_cfpb_final-rule_integrated-mortgage-disclosures.pdf). A useful but skeptical catalogue of CFPB actions, some aimed at manipulative behavior, can be found in Adam Smith and Todd Graziano, *Behavior, Paternalism, and Policy: Evaluating Consumer Financial Protection* (2014), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2408083](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2408083)

manipulation to recent psychological and behavioral findings. Part III turns to normative questions, investigating whether and how manipulation might be unacceptable in light of commitments to autonomy, dignity, and welfare. Part IV examines the relevance of consent, transparency, and democratic authorization. It concludes that while individual consent justifies manipulation, transparency and democratic authorization do not. Part V explores free speech issues and consumer protection. Part VI briefly concludes.

## II. What Manipulation Is

### A. Insulting Deliberation

A great deal of effort has been devoted to the definition of manipulation, almost exclusively within the philosophical literature.<sup>41</sup> Many of the efforts focus on the effects of manipulation in counteracting or undermining people's ability to engage in rational deliberation. On T.M. Wilkinson's account, for example, manipulation "is a kind of influence that bypasses or subverts the target's rational capacities."<sup>42</sup> Wilkinson urges that manipulation "subverts and insults a person's autonomous decision making," in a way that treats its objects as "tools and fools."<sup>43</sup> He thinks that "manipulation is intentionally and successfully influencing someone using methods that pervert choice."<sup>44</sup>

Recall, for example, efforts to enlist attractive people to sell cars, or to use frightening music and ugly photos to attack a political opponent. We might think that in such cases, customers and voters are being insulted in the sense that the relevant speaker is not giving them anything like a straightforward account of the virtues of the car or the vices of the opponent, but is instead using associations of various kinds to press the chooser in the manipulator's preferred direction. On a plausible view, manipulation is involved to the extent that deliberation is insufficient. Here again, it is important to notice that we should speak of degrees of manipulation, rather than a simple on-off switch.

In a related account, Ruth Faden and Tom Beauchamp define psychological manipulation as "any intentional act that successfully influences a person to belief or behavior by causing changes in mental processes *other than those involved in*

---

<sup>41</sup> An excellent overview is *Manipulation* (Christian Coons and Michael Webster eds. 2014).

<sup>42</sup> Christian Coons and Michael Webster, Introduction, in *id.* at 11.

<sup>43</sup> Wilkinson, *supra* note, at 145.

<sup>44</sup> See *id.* An especially valuable discussion, reaching a different conclusion, is Anne Barnhill, What is Manipulation? in *Manipulation: Theory and Practice* 50, 72 (Christian Coons and Michael Weber eds. 2014). Note the emphasis, in defining manipulation, on the manipulator's attempt to influence choices and decisions. If one tries to put people in a certain mood (for example, by taking them to lunch), manipulation is not involved if there is no effort to influence their choices, even when the attempt to lift their mood does not involve an appeal to their reflective and deliberative capacities. (Indeed, the most successful efforts to lift moods often make no such appeal.)

*understanding.*”<sup>45</sup> Joseph Raz suggests that “Manipulation, unlike coercion, does not interfere with a person’s options. Instead it perverts the way that person reaches decisions, forms preferences or adopts goals.”<sup>46</sup>

Of course the idea of “perverting” choice, or people’s way of reaching decisions or forming preferences, is not self-defining; it can be understood to refer to methods that do not appeal to, or produce, the right degree or kind of reflective deliberation. If so, an objection to manipulation is that it “infringes upon the autonomy of the victim by subverting and insulting their decision-making powers.”<sup>47</sup> The objection also offers one account of what is wrong with lies, which attempt to alter behavior not by engaging people on the merits and asking them to decide accordingly, but by enlisting falsehoods, usually in the service of the liar’s goals (an idea that also points the way to a welfarist account of what usually makes lies wrong<sup>48</sup>). A lie is disrespectful to its objects, not least if it attempts to exert influence without asking people to make a deliberate choice in light of relevant facts. But when lies are not involved, and when the underlying actions appear to be manipulative, the challenge is to concretize the ideas of “subverting” and “insulting.”<sup>49</sup>

It is tempting to adopt a simple definition, to this effect: *A statement or action is manipulative to the extent that it does not engage or appeal to people’s capacity for reflective and deliberative choice.* The problem with this definition is that it is far too broad, sweeping up much action that is a standard part of daily life, and that it is rarely taken as manipulative. Suppose, for example, that a good friend frames an option in the most attractive light and with a cheerful voice; or that the Department of Transportation embarks on a vivid, even graphic public education campaign to reduce texting while driving<sup>50</sup>; or that a politician argues in favor of same-sex marriage in a way that points, in an emotionally evocative way, to the lived experience of same-sex couples. In all of these cases, we might have long debates about whether the relevant statements are appealing to people’s capacity for reflective and deliberative choice. And even if we conclude that they are not, we should not therefore be committed to the view that manipulation is involved.

To warrant that conclusion, the word “sufficiently” is required, to add the suggestion that people have been in some sense tricked or fooled, or at least that their deliberative capacities have not been adequately engaged. In this sense, there is a

---

<sup>45</sup> Ruth Faden and Tom Beauchamp, *A History and Theory of Informed Consent* 354-68 (1986).

<sup>46</sup> Joseph Raz, *The Morality of Freedom* 377-79 (1986).

<sup>47</sup> See Wilkinson, *supra* note.

<sup>48</sup> Of course some lies are justified; the intentions of the liar might matter (for example, to spare someone’s feelings), and the consequences might be exculpatory (to prevent serious harm). See Bok, *supra* note.

<sup>49</sup> For relevant discussion in the context of deception, see Craswell, *supra* note.

<sup>50</sup> This is not hypothetical. See <http://www.dot.gov/briefing-room/us-department-transportation-releases-new-“faces-distracted-driving”-video>

connection between the idea of manipulation and the idea of deceit; we can even see the former as a lighter or softer version of the latter. With an act of deceit, people almost inevitably<sup>51</sup> feel betrayed and outraged once they are informed of the truth. The same is true of manipulation. Once the full context is revealed, those who have been manipulated tend to feel used. They ask: *Why wasn't I allowed to decide for myself?*

In an illuminating discussion, with implications for policy and law, Anne Barnhill defines manipulation as “directly influencing someone’s beliefs, desires, or emotions, such that she falls short of ideals for belief, desire, or emotion in ways typically not in her self-interest or likely not in her self-interest in the present context.”<sup>52</sup> Notwithstanding its ambiguity and need for specification, the idea of “falling short of ideals” is helpful, and it should be seen as an effort to capture the same idea as the word “sufficiently.” Note that the standard here is best taken as objective, not subjective. The question is whether someone has, in fact, sufficiently engaged a chooser’s deliberative capacities – not whether the chooser so believes. But there is a problem with Barnhill’s definition, which is that it excludes, from the category of manipulation, influences that are in the self-interest of the chooser. Some acts of manipulation count as such even if they leave the chooser better off. (You might be manipulated to purchase a car that you end up much enjoying.) We might say that such acts are justified – but they are manipulative all the same.

To understand manipulation in this general way, it should not be necessary to make controversial claims about the nature of choice or the role of emotions. We should agree that many decisions are based on unconscious processing and that people often lack a full sense of the wellsprings of their own choices.<sup>53</sup> Even if this is so, a manipulator might impose some kind of influence that unduly undermines or bypasses reflection and deliberation. It is possible to acknowledge the view that emotions might themselves be judgments of value<sup>54</sup> while also emphasizing that manipulators attempt to influence people’s choices without promoting much in the way of reflective thinking about the values at stake. In ordinary language, the idea of manipulation is invoked by people who are not committed to controversial views about psychological or philosophical questions, and it is best to understand that idea in a way that brackets the relevant controversies.

## **B. Manipulating System 1**

We can make progress in understanding some kinds of manipulation with reference to the now-widespread view that the human mind contains not one but two “cognitive

---

<sup>51</sup> They might be grateful, however, if the deceit was genuinely undertaken in order to promote their interests.

<sup>52</sup> Anne Barnhill, What is Manipulation? in *Manipulation: Theory and Practice* 50, 72 (Christian Coons and Michael Weber eds. 2014). Barnhill builds on Robert Noggle, *Manipulative Actions: A Conceptual and Moral Analysis*, 34 *Am Phil Q* 57 (1995).

<sup>53</sup> See Daniel Kahneman, *Thinking, Fast and Slow* (2011).

<sup>54</sup> See Martha Nussbaum, *Upheavals of Thought* (2003).

systems.”<sup>55</sup> In the social science literature, the two systems are described as System 1 and System 2.<sup>56</sup> System 1 is the automatic, intuitive system, prone to biases and to the use of heuristics, while System 2 is more deliberative, calculative, and reflective. Manipulators often target System 1, and they attempt to bypass or undermine System 2.<sup>57</sup>

System 1 works quickly. Much of the time, it is on automatic pilot. It is driven by habits. When it hears a loud noise, it is inclined to run. When it is offended, it wants to hit back. It certainly eats a delicious brownie. It can procrastinate; it can be impulsive. It is easy to manipulate. It wants what it wants when it wants it. It can be excessively fearful and too complacent. It is a doer, not a planner.

By contrast, System 2 is reflective and deliberative. It calculates. It hears a loud noise, and it assesses whether the noise is a cause for concern. It thinks about probability, carefully though sometimes slowly. It does not really get offended. If it sees reasons for offense, it makes a careful assessment of what, all things considered, ought to be done. It sees a delicious brownie, and it makes a judgment about whether, all things considered, it should eat it. It is hardly to manipulate. It insists on the importance of self-control. It is a planner more than a doer.

We need venture contested claims about the nature of the two systems<sup>58</sup> in order to find it helpful to suggest that many actions count as manipulative because they appeal to System 1, and because System 2 is being subverted, tricked, undermined, or insufficiently involved or not informed. Consider the case of subliminal advertising, which should be deemed manipulative, because it operates “behind the back” of the person involved, without appealing to his conscious awareness. People’s decisions are affected in a way that entirely bypasses their own deliberative capacities. If this is the defining problem with subliminal advertising, we can understand why involuntary hypnosis would also count as manipulative. But almost no one favors subliminal advertising, and to say the least, the idea of involuntary hypnosis lacks much appeal. The question is whether admittedly taboo practices can shed light on actions that are more familiar or that might be able to command broader support.

#### **D. Illustrative (and Not Easy) Cases**

---

<sup>55</sup> Kahneman, *supra* note; Timothy Wilson, *Strangers to Ourselves* (2014).

<sup>56</sup> *See* Kahneman, *supra* note.

<sup>57</sup> For related observations, see Hansen and Jespersen, *supra* note.

<sup>58</sup> At this point, it might be asked: What, exactly, are these systems? The best answer is that the idea of two systems is a heuristic device, a simplification that is designed to refer to automatic, effortless processing and more complex, effortful processing. But it is also true that identifiable regions of the brain are active in different tasks, and hence it may well be right to suggest that the idea of “systems” has physical referents. An influential discussion states that “[a]utomatic and controlled processes can be roughly distinguished by where they occur in the brain.” Colin Camerer et al., *Neuroeconomics: How Neuroscience Can Inform Economics*, 43 *J. ECON. LITERATURE* 1, 17 (2005).

Consider some cases that test the boundaries of the concept of manipulation.

1. Suppose that public officials try to persuade people to engage in certain behavior with the help of relative risk information: “If you do not do X, your chances of death from heart disease will triple!”<sup>59</sup> But suppose that for the relevant population, the chance of death from heart disease is very small – say, one in 100,000 – and that people are far more influenced by the idea of “tripling the risk” than they would be if they learned that if they do not do X, they could increase a 1/100,000 risk to a 3/100,000 risk (to say the least, a modest increase). The relative risk frame is far more attention-grabbing than the absolute risk frame; a tripling of a risk sounds alarming, but if the increase is by merely 2/100,000, people might not be much concerned. It is certainly reasonable to take the choice of the relative risk frame (which suggests a large impact on health) is an effort to frighten people and thus to manipulate them (at least in a mild sense).

It is true that any description of a risk requires some choices; people who describe risks cannot avoid some kind of framing. But framing is not the same as manipulation. There is a good argument this particular choice does not sufficiently engage, or show a great deal of respect for, people’s deliberative capacities; it might even be an effort to aim specifically at System 1. As we shall see, that conclusion does not mean that the use of the relative risk frame is necessarily out of bounds.<sup>60</sup> This is hardly the most egregious case of manipulation, and if it saves a number of lives across a large population, it might be justified. But it can be counted as manipulative.

2. Suppose that public officials are alert to the power of loss aversion,<sup>61</sup> and hence they use the “loss frame,” so as to trigger people’s concern about the risks associated with obesity and excessive energy consumption. They might deliberately choose to emphasize, in some kind of information campaign, how much people would *lose from not using* energy conservation techniques, rather than how much people would *gain from using* such techniques.<sup>62</sup> Is the use of loss aversion, with its predictably large effects,<sup>63</sup> a form of manipulation?

The answer is not obvious, but there is a good argument that it is not, because deliberative capacities remain sufficiently involved. Even with a loss frame, people remain fully capable of assessing overall effects. But it must be acknowledged that the

---

<sup>59</sup> Wilkinson, *supra* note, at 347, uses this example.

<sup>60</sup> See Craswell, *Regulating Deceptive Advertising*, *supra* note, at 552: “Advertisements are potentially dangerous ‘products,’ and advertisers should take reasonable steps to prevent consumers from being harmed by their products. But advertisers should not be faulted if any further changes would have made matters worse for consumers rather than better.”

<sup>61</sup> See Eyal Zamir, *Law, Psychology, and Morality: The Role of Loss Aversion* (2014).

<sup>62</sup> See Elliott Aronson, *The Social Animal* 124-25 (6th ed. 1996).

<sup>63</sup> See *id.*

deliberate use of loss aversion might be an effort to trigger the negative feelings that are distinctly associated with losses.

Here too, it is a separate question whether the use of loss aversion raises serious ethical objections. Within the universe of arguably manipulative statements, those that enlist loss aversion hardly count as the most troublesome, and in the case under discussion, the government's objectives are laudable. If the use of loss aversion produces large gains (in terms of health or economic benefits), we would not have much ground for objection.<sup>64</sup> But we can identify cases in which the use of loss aversion is self-interested, and in which the surrounding context makes it a genuine example of manipulation.<sup>65</sup> Consider, for example, the efforts of banks, in the aftermath of a new regulation from the Federal Reserve Board, to enlist loss aversion to encourage customers to opt-in to costly overdraft protection programs by saying, "*Don't lose your ATM and Debit Card Overdraft Protection*" and "*STAY PROTECTED with [ ] ATM and Debit Card Overdraft Coverage*."<sup>66</sup> In such cases, there is an evident effort to trigger a degree of alarm, and hence it is reasonable to claim that customers were being manipulated, and to their detriment.

3. Alert to the behavioral science on social influences,<sup>67</sup> a planner might consider the following approaches:

- a. Inform people that most people in their community *are engaging in undesirable behavior* (drug use, alcohol abuse, delinquent payment of taxes, environmentally harmful acts).
- b. Inform people that most/many people in their community *are engaging in desirable behavior*.
- c. Inform people that most/many people in their community *believe that people should engage in certain behavior*.

The first two approaches rely on "descriptive norms," that is, what people actually do.<sup>68</sup> The second approach relies on "injunctive norms," that is, what people think that people should do. As an empirical matter, it turns out that descriptive norms are ordinarily more powerful.<sup>69</sup> If a change in behavior is what is sought, then it is best to emphasize that most/many people actually do the right thing.<sup>70</sup> But if most/many people do the wrong thing, it can be helpful to invoke injunctive norms.<sup>71</sup>

---

<sup>64</sup> Craswell, *Regulating Deceptive Advertising*, supra note, at 551-555.

<sup>65</sup> See Lauren E. Willis, *When Defaults Fail: Slippery Defaults*, 80 U. Chi. L. Rev. 1155 (2012).

<sup>66</sup> *Id.* at 1192.

<sup>67</sup> For a summary, see Richard H. Thaler and Cass R. Sunstein, *Nudge* (2008).

<sup>68</sup> See Robert Cialdini, *Crafting Normative Messages to Protect the Environment*, 12 *Current Directions in Psychological Science* 105 (2003).

<sup>69</sup> *Id.*

<sup>70</sup> *Id.*

<sup>71</sup> *Id.*; Wesley Schultz et al., *The Constructive, Destructive, and Reconstructive Power of*



Suppose that a public official is keenly aware of these findings and uses them. Is he engage in manipulation? The word “sufficiently” becomes relevant here as well. Without doing much violence to ordinary language, some people might think it reasonable to conclude that it is manipulative to choose the formulation that will have the largest impact. At least this is so if social influences work as they do because of their impact on the automatic system, and if they bypass deliberative processing.<sup>72</sup> But as an empirical matter, this is far from clear; information about what other people do, or what other people think, can be part of reflective deliberation, and hardly opposed to it. So long as the official is being truthful, it would strain the boundaries of the concept to accuse him of manipulation: When they are informed about what most people do, people’s powers of deliberation are sufficiently engaged.<sup>73</sup>

4. Default rules often stick, in part because of the force of inertia, in part because of the power of suggestion.<sup>74</sup> Suppose that a public official is aware of that fact and decides to reconsider a series of default rules in order to exploit the stickiness of defaults. Seeking to save money, she might decide in favor of a double-sided default for printers.<sup>75</sup> Seeking to reduce pollution, she might promote, or even require, a default rule in favor of green energy.<sup>76</sup> Seeking to increase savings, she might promote, or even require, automatic enrollment in retirement plans.<sup>77</sup>

Are these initiatives manipulative? Insofar as the default rules carries an element of suggestion -- a kind of informational signal -- they are not. Such rules appeal to deliberative capacities insofar as they convey information about what planners think people ought to be buying. The analysis is less straightforward insofar as default rules stick because of inertia: Without making a conscious choice, people end up enrolled on some kind of program or plan. In a sense, the official is exploiting System 1, which is

---

Social Norms, 18 Psych Science 429 (2007).

<sup>72</sup> For relevant (but not decisive) findings, see Caroline J. Charpentier et al., The Brain’s Temporal Dynamics from a Collective Decision to Individual Action, 34 J. Neuroscience 5816 (2014).

<sup>73</sup> Note here the related finding that notwithstanding the power of suggestion, and accompanying social influences, people reject default rules that they do not like. See Zachary Brown et al., *Testing the Effects of Defaults on the Thermostat Settings of OECD Employees*, 39 Energy Econ. 128 (2013).

<sup>74</sup> See Eric Johnson and Daniel Goldstein, Decisions by Default, in *The Behavioral Foundations of Public Policy* 417 (Eldar Shafir ed. 2013); Cass R. Sunstein, *Choosing Not To Choose* (2015).

<sup>75</sup> See Johan Egebark and Mathias Ekström, *Can Indifference Make the World Greener?* (2013), available at [http://www2.ne.su.se/paper/wp13\\_12.pdf](http://www2.ne.su.se/paper/wp13_12.pdf)

<sup>76</sup> See Cass R. Sunstein and Lucia Reisch, *Automatically Green: Behavioral Economics and Environmental Protection*, 38 Harv Env L Rev 127 (2014).

<sup>77</sup> Shlomo Benartzi & Richard H. Thaler, *Behavioral Economics and the Retirement Savings Crisis*, 339 Science 1152 (2013).

prone to inertia and procrastination.<sup>78</sup> The question is whether automatic enrollment fails “sufficiently” to engage reflection and deliberation.

In answering that question, it is surely relevant that an opt-in default is likely to stick as well, and for the same reasons, which means that the question is whether *any* default rule counts as a form of manipulation. The answer to that question is plain: Life cannot be navigated without default rules, and so long as the official is not hiding or suppressing anything, the choice of one or another should not be characterized as manipulative. Note that people do reject default rules that they genuinely dislike, so long as opt-out is easy – an empirical point in favor of the conclusion that such rules should not be counted as manipulative.<sup>79</sup>

5. *A potpourri*. There is no question that much of modern advertising is directed at System 1, with attractive people, bold colors, and distinctive aesthetics. (Consider advertisements for Viagra.) Often the goal is to trigger a distinctive affect and more specifically to enlist the “affect heuristic,” which puts the question of manipulation in stark relief.<sup>80</sup> Much of website design is an effort to trigger attention and to put it in the right places.<sup>81</sup> Cell phone companies, restaurants, and clothing stores use music and colors in a way that is designed to “frame” products in a distinctive manner. Doctors, friends, and family members (including spouses) sometimes do something quite similar. Is romance an exercise in manipulation? Some of the time, the answer is surely yes, though the question of “sufficiently” raises special challenges in that context.<sup>82</sup>

Acting as advocates, lawyers may be engaged in manipulation; that is part of their job, certainly in front of a jury. The same can be said about some aspects of the provision of medical care, when doctors want patients to choose particular options, and enlist behaviorally informed technique to manipulate them to do so. Or consider certain uses of social media – as, for example, when Facebook attempted to affect (manipulate) the emotions of 689,003 people through the display of positive or negative stories to see how they affected their moods.<sup>83</sup> A great deal of conduct, however familiar, can be counted as manipulative in the relevant sense.

---

<sup>78</sup> See Riccardo Rebonato, *Taking Liberties: A Critique of Libertarian Paternalism* 3-7 (2011).

<sup>79</sup> See Brown, *supra* note; John Beshears et al., *The Limitations of Defaults* (Sept. 15, 2010) (unpublished manuscript), *available at* <http://www.nber.org/programs/ag/rrc/NB10-02,%20Beshears,%20Choi,%20Laibson,%20Madrian.pdf>.

<sup>80</sup> See *THE FEELING OF RISK: NEW PERSPECTIVES ON RISK PERCEPTION* 3-20 (Paul Slovic ed., 2010).

<sup>81</sup> Steve Krug, *Don't Make Me Think Revisited: A Common Sense Approach to Web and Mobile Usability* (2014).

<sup>82</sup> I am aware of no detailed treatment of this question, but for relevant discussion, see Eric Cave, *Unsavory Seduction and Manipulation*, in *Manipulation*, *supra* note, at 176.

<sup>83</sup> See Adam Kramer et al., *Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks*, 111 *PNAS* 8788 (2014).

### III. What Is Wrong With Manipulation, So Understood?

#### A. Autonomy and Dignity

1. *Respect.* The most obvious problem with manipulation is that it can insult both autonomy and dignity. From the standpoint of autonomy, the problem is that manipulation can deprive people of agency, resting on a continuum for which coercion is the endpoint. (If people are manipulated into buying a product, they might feel coerced.) From the standpoint of dignity, the problem is that manipulation can be humiliating. Healthy adults, not suffering from a lack of capacity, should not be tricked; they should be treated as fully capable of making their own decisions. Their authority over their own lives should not be undermined by approaches that treat them as children or as puppets. An act of manipulation does not treat people with respect.<sup>84</sup>

Suppose, for example, that someone thinks, “I want all my friends to do certain things, and I know a number of strategies to get them to do those things. I have read a great deal of psychology and behavioral science, including the best work on social influence, and my project is to use what I know to manipulate my friends.”<sup>85</sup> Such a person would not be respecting her friends’ autonomy. She would be using them as her instruments. Indeed, her actions would be inconsistent with the nature of friendship itself, which entails a relationship that is not strictly or mostly instrumental.

Now turn to the case of government. Suppose that public officials – say, in a governor’s office – similarly learn a great deal about how to influence people, and suppose that they decide to use what they learn to achieve certain policy goals. Suppose that some of the relevant instruments attempt to subvert or bypass deliberation. We need to know the details – what, exactly, are they doing? -- but it could well be fair to say that manipulation is involved, and that public officials are not sufficiently respecting their citizens’ autonomy.<sup>86</sup> Again we need to know the details, but it could also be fair to say that such officials are not treating citizens with respect; they might be using them as

---

<sup>84</sup> See Marcia Baron, *Manipulativeness*, 77 *Proceedings and Addresses of the American Philosophical Association* 37 (2003), and in particular this suggestion: “By contrast, the person who has the virtue corresponding to manipulateness – a virtue for which we do not, I believe, have a name – knows when it is appropriate to try to bring about a change in another’s conduct and does this for the right reasons, for the right ends, and only where it is warranted (and worth the risks) and only using acceptable means. The virtuous person tries to reason with the other, not cajole or trick him into acting differently. . . . being manipulative is a vice because of its arrogance and presumption, and because the manipulative person is too quick to resort to ruses . . . .” *Id.* at 48, 50.

<sup>85</sup> A classic source here is Robert Cialdini, *Influence* (2006); another classic, far less academic but highly informative, is Dale Carnegie, *How to Win Friends and Influence People* (1998). Both books can be seen as offering a great deal of advice about successful manipulation.

<sup>86</sup> See Sunstein, *Ethics and Choice Architecture*, *supra* note, for detailed discussion.

instruments or as puppets for their own ends (or for the perhaps public-spirited ends that they favor).<sup>87</sup>

2. *Role*. We should be able to see, in this light, that *role* greatly matters to the assessment of manipulation. Suppose that Jones is trying to obtain a job. It is hardly unacceptable for Jones to attempt to get prospective employers to like him, and if Jones learns about social influence and the operations of System 1, it would hardly be illegitimate for him to take advantage of what he learns. To be sure, there are ethical limits on what Jones can do, and even for someone seeking a job, the most egregious forms of manipulation would cross the line. But in interactions or relationships that are instrumental, and that are so understood, deontological constraints are weakened or at least different.

In an advertising campaign, everyone knows the nature of the interaction. Manipulation is the coin of the realm. The purpose of advertisements is to sell products, and while we can find purely factual presentations, many advertisements do not appeal to reflection or deliberation at all. They try to create certain moods and associations. Something similar can be said about some aspects of political campaigns. The relationship between a campaign and voters has an instrumental character: Campaigns want votes, and everyone understands that. In the process, both advertisements and speeches will have manipulative features. It would be extravagant to say that in such cases, people have consented to manipulation in all imaginable forms. Here too, lines can be crossed.<sup>88</sup> But it is important that people are aware of the distinctive nature of the relevant enterprises.

Other roles are accompanied by different norms, and manipulation might not fit with, or might even violate, those norms. When governments deal with their citizens, they face radically different norms from those that apply in campaigns. At least this is so in free and democratic societies, in which it is understood that the public is ultimately sovereign. To be sure, public officials are hardly forbidden from framing options in a way that casts those they prefer in the most favorable light. But as the manipulative characteristics of their actions become more extreme, the scope for legitimate objection becomes greater.<sup>89</sup>

## B. Welfare

Suppose that we are welfarists, and that we believe that what matters is how people's lives are going.<sup>90</sup> Suppose too that we care about violations of autonomy and

---

<sup>87</sup> Nudges do not run afoul of this conclusion, at least most of the time. See *id.*

<sup>88</sup> See Oren Bar-Gill, *Seduction By Contract* (2012), for a superb discussion.

<sup>89</sup> See Goodin, *supra* note.

<sup>90</sup> I am bracketing here various questions about how welfarism is best understood. It is possible to have a conception of welfare that includes consideration of autonomy and dignity. See Amartya Sen, *Development As Freedom* (2000); Martha Nussbaum,

dignity only insofar as such violations affect people's subjective experiences (for example, by making them feel confined or humiliated). If, how should we think about manipulation?

It should be clear that there is no taboo on the practice. As we shall see, manipulation might promote people's welfare. But there is a distinctive welfarist objection to manipulation, which takes the following form. As a general rule, choosers know what is in their best interest (at least if they are adults, and if they do not suffer from a problem of capacity<sup>91</sup>). They have unique access to their situations, their constraints, and their tastes.<sup>92</sup> If they are manipulated, they are deprived of the (full) ability to make choices on their own, simply because they are not give a fair or adequate chance to weigh all variables. If someone wants to help people to make better choices, his obligation is to inform them, so that they can themselves engage in such weighing.

The problem with the manipulator is that he lacks relevant knowledge – about the chooser's situation, tastes, and values.<sup>93</sup> Lacking that knowledge, he nonetheless subverts the process by which choosers make their own decisions about what is best for them. Things are even worse if the manipulator is focused on his own interests rather than on those of choosers. It is in this sense that a self-interested manipulator can be said to be stealing from people – both limiting their agency and moving their resources in the preferred direction.

For these reasons, the welfarist objection to paternalism is rooted in the same concerns that underlie Mill's Harm Principle. In Mill's view, the problem with outsiders, including government officials, is that they lack the necessary information. Mill insists that the individual "is the person most interested in his own well-being,"<sup>94</sup> and the "ordinary man or woman has means of knowledge immeasurably surpassing those that can be possessed by any one else."<sup>95</sup> When society seeks to overrule the individual's judgment, it does so on the basis of "general presumptions," and these "may be altogether wrong, and even if right, are as likely as not to be misapplied to individual cases."<sup>96</sup>

---

Creating Capabilities (2011); Utilitarianism and Welfarism, 76 J Phil 463 (1979). For instructive discussion, see Matthew Adler, *Welfare and Fair Distribution* (2011).

<sup>91</sup> A child, or a person suffering from some form of dementia, has a weaker objection to manipulation. Parents manipulate young children all the time, partly to promote their welfare. Caretakers manipulate people who are suffering from dementia. These practices are largely taken for granted, but we could imagine situations in which they would raise serious ethical questions. Even if the relevant manipulation is in the interest of those who are being manipulated, the interests in autonomy and dignity impose constraints even here.

<sup>92</sup> See Hayek, *supra* note.

<sup>93</sup> *Id.*

<sup>94</sup> Mill, *supra* note.

<sup>95</sup> *Id.*

<sup>96</sup> *Id.*

These points apply to those engaged in manipulation no less than to those engaged in coercion.

Notwithstanding these points, it should be clear that from the welfarist standpoint, there should be no ban on manipulation.<sup>97</sup> Everything depends on whether manipulation improves people's welfare. To see the point, imagine a benign, all-knowing, welfare-promoting manipulator – a kind of idealized parent – who is concerned only with the welfare of those who are being manipulated, who has all the knowledge he needs, and who simply does not make mistakes. By hypothesis, the welfare-promoting manipulator should be celebrated on welfarist grounds. The major qualification is that if people know that they are being manipulated, and do not like it, there will be a welfare loss, and that loss will have to be counted in the overall assessment. The simple point is that if people hate manipulators, manipulation is less likely to be supportable on welfare grounds (unless it is hidden, which raises problems of its own<sup>98</sup>).

But put that point to one side. The main problem with the thought experiment is that manipulators are unlikely to be either benign or all-knowing. Often they have their own agendas, and the fact that they engage in manipulation attests to that fact. If they are genuinely concerned about the welfare of the chooser, why not try to persuade them? Why cross the line to manipulation? To be sure, the manipulator might be able to answer this question if, for example, time is of the essence, or if the chooser lacks capacity (because, for example, he is a child or very ill). Or suppose that graphic health warnings, aimed directly at System 1, save numerous lives; suppose too that numerous lives cannot be saved with a merely factual presentation unaccompanied by graphic health warnings. On welfarist grounds, a great deal might be said on behalf of graphic health warnings.<sup>99</sup>

The example shows that from the standpoint of welfare, everything depends on the context; the fact that manipulation is involved does not *necessarily* impeach the manipulator's welfare calculus. But in many situations, suspicion about manipulators' goals is perfectly justified. To this point it must be added that even when those goals are admirable, manipulators may not know enough to justify their actions. Consider Friedrich Hayek's remarkable suggestion that "the awareness of our irremediable ignorance of

---

<sup>97</sup> Note that the question here is whether manipulation increases or decreases welfare; it is not whether the law, or some regulator, should ban manipulation. The latter question raises issues about institutional competence and decision costs. It also requires attention to the effect of manipulation on large populations with heterogeneous understandings. In response to an advertising campaign, for example, some people might be manipulated (in the sense that System 1 is essentially all that is affected) while others are not (because the campaign triggers a significant amount of deliberation).

<sup>98</sup> Hidden manipulation is risky, because it might be disclosed, and people will not be happy to learn that it has been hidden.

<sup>99</sup> Christine Jolls, Product Warnings, Debiasing, and Free Speech: The Case of Tobacco Regulation, 169 *J Institutional and Theoretical Economics* 53 (2013).

most of what is known to somebody [who is a planner] is *the chief basis of the argument for liberty*.”<sup>100</sup> I will return to this point, and the question of regulation, in Part VI.

## V. Consent, Transparency, Democracy

### A. Manipulation With Consent: “I Welcome It!”

Suppose that people consent to manipulation.<sup>101</sup> An alcoholic might tell his wife: “I am trying so hard to quit. Please use whatever techniques you can think of to help me. Manipulation is very much on the table. I welcome it!” Or suppose that the overwhelming majority of smokers tell their government: “I want to stop! If you can find a way to help me to overcome my addiction, I would be grateful.” T. M. Wilkinson notes that it is too crude to say that manipulation infringes upon autonomy, because “manipulation could be consented to. If it were consented to, in the right kind of way, then the manipulation would at least be consistent with autonomy and might count as enhancing it.”<sup>102</sup>

The conclusion has a great deal of force. We can understand consent as suggesting support from System 2, which might welcome a little manipulation (or possibly a lot) as a way of cabining the adverse effects of System 1. The tale of Ulysses and the Sirens is instructive here, whether Ulysses was requesting manipulation or something else.<sup>103</sup> Nor is there an objection, in the case of consent, from the standpoint of welfare. The chooser has decided that he will be better off if he is manipulated. If we see his choice as presumptively promoting his welfare, we should respect it, even if what he chose is manipulation.

In the easiest cases, consent is explicit. In harder cases, it is only implicit, in the sense that the manipulator infers it from the circumstances or believes, with good reason, that the chooser would consent if asked. If the inference is less than reliable, the consent justification is correspondingly weakened. If the belief is reasonable but potentially wrong, it might make sense to insist on obtaining explicit consent in order to avoid the risk of error. It is important to see that consensual manipulation is an unusual case; those who need help do not ordinarily ask, “Please manipulate me.” But such cases certainly do exist, at least when people face serious problems of self-control.

### B. Transparency and Manipulation

---

<sup>100</sup> See Hayek, *The Market and Other Orders*, supra note (emphasis added).

<sup>101</sup> On one view, the concept of manipulation presupposes a lack of consent. See Robert Goodin, *Manipulatory Politics* 9 (1980) (discussing the idea of “unknown interference”). But the examples given in text suggest that manipulation can be a product of consent and even invitation.

<sup>102</sup> Wilkinson, supra note, at 345.

<sup>103</sup> See Jon Elster, *Ulysses and the Sirens* (1983).

The idea of manipulation is sometimes taken to imply a lack of transparency, as if something important is being hidden or not being disclosed.<sup>104</sup> If a manipulator is acting as a puppeteer, he might be failing to reveal his own role; that can be an important feature of manipulation. With respect to manipulation, however, it is not entirely clear what transparency even means.<sup>105</sup>

Sometimes manipulation itself *consists* in a lack of transparency about a relevant feature of a situation; that is the very manipulation involved. (A parent tells a small child: If you are very good, Santa Claus will bring you a toy giraffe.) Once the relevant feature is revealed, the manipulation is gone. But the manipulative character of some acts does not consist in their hidden quality. (Recall the use of relative risk reduction.) Some acts can be both manipulative and fully revealed to those who are being manipulated.<sup>106</sup> A graphic health warning, for example, is perfectly transparent (and if it is required by regulation, it is even likely to be preceded by a period for public comment<sup>107</sup>). Subliminal advertising could be preceded by an explicit warning: “This movie contains subliminal advertising.”

In the pivotal scene in *The Wizard of Oz*, the Wizard says, “Pay no attention to the man behind the curtain.” The man behind the curtain is of course a mere human being who is masquerading as the great Wizard – and who is both claiming far more authority than he deserves and designing social situations in a way that hides features that, if revealed, would greatly alter people’s judgments and choices. Or consider a less celebrated movie, *The Truman Show*, in which the life course of Truman, the movie’s hero-protagonist, is affected by multiple decisions from a master manipulator, who conceals the facts that Truman is the unwitting star of a television show and that his friends and acquaintances consist of actors. Covertneess and hiding are common features of manipulation. Whenever people who are imposing influence conceal their own role, it seems reasonable to object. A lack of transparency offends both autonomy and dignity.

From the standpoint of welfare, we might ask why, exactly, someone has failed to be upfront with the chooser, who ought to be able to make his own decisions, armed with relevant information. As before, however, the analysis of welfare is more complex, for we could imagine cases in which transparency is not necessary and may in fact be a

---

<sup>104</sup> See Goodin, *supra* note, at 7-12.

<sup>105</sup> On nudging, transparency, and manipulation, see Hansen and Jespersen, *supra* note.

<sup>106</sup> See George Loewenstein et al., *Warning: You Are About To Be Nudged* (2014) (unpublished manuscript). For relevant findings, see Gidon Felsen et al., *Decisional Enhancement and Autonomy: Public Attitudes Toward Overt and Covert Nudges*, 8 *Judgment and Decision Making* 203 (2012). On the fact that manipulation can be transparent rather than covert, see Barnhill, *supra* note: “I think that deceptiveness or covertneess is a favorite technique of manipulators – manipulation is more likely to succeed if its target doesn’t realize what’s happening. But manipulation needn’t be covert. Covertneess isn’t what’s definitive of manipulation.”

<sup>107</sup> As was the case for the FDA regulation invalidated by the court of appeals. See note *supra*.



problem. Suppose that someone suffers from a serious self-control problem and that his life is in danger (from, say, alcohol or drug addiction). Suppose too that a manipulator has come across a life-saving strategy and that transparency would render the manipulation less effective. By hypothesis, welfarist considerations argue against transparency. Points of this kind have the strongest intuitive force when applied to people who lack capacity (young children, the mentally ill), but we can imagine contexts in which adults with full capacity would benefit from being manipulated.<sup>108</sup>

There might also be a welfarist justification for hidden manipulation in other extreme circumstances – as, for example, when people are trying to stop a kidnapping or to save a kidnapping victim. If the goal is to stop a wrongdoer, or someone who threatens to do real harm, it may be perfectly acceptable or even obligatory to manipulate them and to hide that fact. They have forfeited their right to be treated with respect, and their welfare, as choosers, is not a matter of much concern.

In standard cases, however, this argument will be unavailable. It follows that most of the time, manipulation should not be hidden or covert, even when it is justified; return to the case of graphic health warnings. Transparency is a necessary condition. Note, however, that it is not sufficient.<sup>109</sup> Subliminal advertising would not become acceptable merely because people were informed about it. If a movie chain announced that its previews would be filled with subliminal advertisements, people could fairly object.

### **C. Democratically Authorized Manipulation**

What if manipulation is democratically authorized? Suppose that a national legislature expressly votes for it, perhaps in order to improve public health (as, for example, by discouraging smoking or unhealthy eating), or perhaps to promote other goals (such as enlistment in the military or adoption of the currently preferred ideology). In relatively benign cases, involving little or no manipulation, a legislature might support an educational campaign that is designed to reduce illnesses and deaths and that enlists a series of behaviorally informed strategies, targeting System 1, to accomplish its goals.

It should be clear that democratic authorization ought not by itself to dissolve otherwise reasonable objections to manipulation. The most obvious problems arise if the national legislature has illegitimate ends (say, the inculcation of racial prejudice or self-entrenchment of some kind). But the familiar objections – involving autonomy, dignity, and welfare -- apply even if the ends are legitimate. If a national legislature authorizes subliminal advertising, it remains fully possible to object on grounds of both autonomy and dignity. An objection from the standpoint of welfare is also possible: Why did the

---

<sup>108</sup> Sarah Conly, *Against Autonomy* 149-72 (2012), offers a series of cases in which, on her view, coercion is justified, among other things on welfare grounds. If manipulation passed the relevant tests (involving cost-benefit analysis and cost-effectiveness, *id.* at 150-52), it would be justified on similar grounds. Conly herself offers cautionary notes about manipulation, pointing to its uneasy relationship with autonomy. *Id.* at 30-31.

<sup>109</sup> See Hansen and Jespersen, *supra* note.

democratic process authorize manipulation, rather than some other kind of communication?

To be sure, we could understand democratic authorization as a form of majority or collective consent, suggesting support from System 2, which might welcome a little manipulation (or possibly a lot) as a way of cabining the adverse effects of System 1. In general, however, there are evident risks in authorizing public officials to pursue this line of argument. The objection to manipulation comes from individuals, who do not want to be manipulated; the fact that a majority wants to manipulate them is no defense.

But in certain contexts, the argument on behalf of at least a modest degree of manipulation might not be implausible, and it is strengthened if the democratic process has supported it. Imagine, for example, a public education campaign that is designed to reduce the risks associated with texting while driving,<sup>110</sup> or an effort to combat the use of dangerous drugs or to convince people to stay in school. Many such campaigns are vivid and have an emotional component; they can be understood as efforts to combat self-control problems and to focus people on the long term.<sup>111</sup>

If government is targeting System 1 – perhaps through framing, perhaps through emotionally evocative appeals – it might be responding to the fact that System 1 has already been targeted, and to people’s detriment. In the context of cigarettes, for example, it is plausible to say that previous manipulations – including advertising and social norms – have influenced people to become smokers. If this is so, perhaps we can say that public officials are permitted to meet fire with fire. But some people might insist that two wrongs do not make a right – and that if the government seeks to lead people to quit, it must treat them as adults, and appeal to their deliberative capacities.

Recall that there are degrees of manipulation, and there is a large difference between a lie and an effort to frame an alternative in an appealing, unappealing, or ugly light. In ordinary life, we would not be likely to accuse our friends or loved ones of manipulation if they characterized one approach as favored by most members of our peer group, or if they emphasized the losses that might accompany an alternative that they abhor, or if they accompanied a description of one option with a grave look and a frown. These are at most mild forms of manipulation, to be sure, and it is important to see that mild forms might well be acceptable and benign (and a bit fun) if they promote the interests of those people at whom they are aimed. No legal system has a general tort called “exploitation of cognitive biases.”

#### **D. Unifying Strands**

If the various arguments are put together, we might be able to evaluate acts of manipulation with the help of the following matrix:

---

<sup>110</sup> See note supra.

<sup>111</sup> To be sure, educational campaigns of this kind might not be counted as manipulation at all, because deliberative capacities are sufficiently engaged.

	<b>Benign and informed</b>	<b>Malign or uninformed</b>
<b>Not highly manipulative</b>	Acceptable on welfare grounds; might be acceptable by reference to autonomy or dignity	Unacceptable
<b>Highly manipulative</b>	Acceptable on welfare grounds; objectionable on grounds of autonomy and dignity	Highly unacceptable

The matrix helps to orient the appropriate response to manipulation from the standpoint of ethics, politics, and law, and indeed it captures widespread intuitions. In the bottom right cell, we can find actions by self-interested or venal public officials both in undemocratic systems<sup>112</sup> and occasionally in democratic ones as well.<sup>113</sup> In the top right, we can find foolish or venal statements or actions by private and public actors that do not entirely bypass people’s deliberative capacities, but that hardly do justice to them. Many government agencies, and many ordinary companies, act in accordance with the top left cell; they portray their behavior in an appealing light, and they try to attract favorable attention, but the particular form of manipulation is hardly egregious. Some governments, some of the time, act in a way that fits in the bottom left cell, perhaps especially with graphic campaigns.

The matrix also provides a start toward an analysis of how the legal system should respond to manipulation. From the welfarist standpoint, the central question is whether the benefits of restricting the manipulative action or statement justify the costs. To answer that question, we need to know what would happen if the action or statement were not made (or were transformed into a non-manipulative version). In this respect, the analysis of manipulation closely parallels the analysis of deception.<sup>114</sup> The costs of manipulation depend, in large part, on whether the manipulator is malign or uninformed. To the extent that it is, there is a risk of serious welfare losses. But suppose that an advertiser is part of a well-functioning competitive process, and that its advertisement includes a degree of manipulation in order to sell a product. If the competitive process is genuinely well-functioning, consumers are not likely to lose much, and market pressures will discipline the use and the effectiveness of manipulation.<sup>115</sup>

The question, then, is whether some kind of market failure exists, so that manipulative behavior can persist or be rewarded. In light of information problems and behavioral biases, the answer is likely to be affirmative, at least in some markets.<sup>116</sup> Of course consumers will have diverse understandings of, and reactions to, statements and

---

<sup>112</sup> See note supra.

<sup>113</sup> See Goodin, supra note, for many examples.

<sup>114</sup> Craswell, *Interpreting Deceptive Advertising*, supra note.

<sup>115</sup> See Edward L. Glaeser, *Paternalism and Psychology*, 73 U. Chi. L. Rev. 133 (2006).

<sup>116</sup> See Bar-Gill, supra note.

actions that plausibly fall within the category of manipulation.<sup>117</sup> Empirical testing of representative populations could provide highly informative here. The fact of heterogeneous understandings will create serious challenges for regulators seeking to prevent arguably harmful forms of manipulation. As we shall see, however, some manipulative acts are so plainly welfare-reducing that it makes sense to restrict them.

## VI. Manipulation, Freedom of Speech, and Consumer Protection

Under established doctrine, government can regulate threats<sup>118</sup>; it can also regulate false or misleading commercial speech<sup>119</sup> and certain forms of coercive speech.<sup>120</sup> May it also regulate manipulation? It is also clear that government can compel certain kinds of speech.<sup>121</sup> May it compel speech that is arguably manipulative?

### A. Compelled Speech

As a testing case, consider the efforts of the Food and Drug Administration (FDA) to require cigarette packages to contain graphic health warnings. The court of appeals invalidated the requirement on the first amendment grounds, concluding that the FDA lacked sufficient evidence to justify the compelled speech.<sup>122</sup> In so ruling, the court did not emphasize the arguably manipulative nature of the graphic warnings. But the lower court opinion did exactly that.<sup>123</sup>

The court found it relevant that “the Rule's graphic-image requirements are *not* the type of purely factual and uncontroversial disclosures that are reviewable under this less stringent standard.”<sup>124</sup> It added, plausibly, that “it is abundantly clear from viewing these images that the emotional response they were crafted to induce is calculated to provoke the viewer to quit, or never to start, smoking: an objective wholly apart from disseminating purely factual and uncontroversial information.”<sup>125</sup> The court concluded that when the government compels speech that does not involve the “purely factual and uncontroversial,” it has to meet a higher burden of justification.

The central idea here lacks support in Supreme Court decisions, but it has some appeal: The first amendment imposes particular barriers to government efforts to require speech that does not merely appeal to deliberative or reflective capacities, but that

---

<sup>117</sup> See Craswell, *Interpreting Deceptive Advertising*, *supra* note, at 672-75.

<sup>118</sup> *Watts v. United States*, 394 US 705 (1969).

<sup>119</sup> *Virginia State Pharmacy Board v. Virginia Citizens Consumer Council*, 425 US 748 (1976).

<sup>120</sup> See note *supra*.

<sup>121</sup> See Note, *The Future of Government-Mandated Health Warnings*, 163 U. Pa. L. Rev. 177 (2014).

<sup>122</sup> See note *supra*.

<sup>123</sup> See *RJ Reynolds v. FDA*, F. Supp. (2011).

<sup>124</sup> *Id.* at

<sup>125</sup> *Id.* at

engages and attempts to activate System 1. On this view, there is no firm rule against compelling manipulative speech of that kind (so long as it is not false or deceptive), but if government is engaging in such compulsion, it must have a strong justification for doing so.

This analysis raises an assortment of issues. Do the graphic warnings count as manipulative? They are certainly designed to create a visceral response (and they do exactly that). But the question is whether they do not sufficiently engage or appeal to people's capacity for reflective and deliberative choice. Here the answer has to come from specifying the idea of "sufficiently." There is an empirical component to the specification: What, exactly, do people understand after they see the warnings? Suppose that for a large part of the population, understanding is actually improved. If so, there is a good argument that manipulation is not involved.<sup>126</sup> But suppose that understanding is not improved. Can the warnings nonetheless be justified?

We could imagine two such justifications. The first is welfarist: Graphic health warnings will save a significant number of lives, and purely factual information will have a far weaker effect. If this is so, then the graphic health warnings do have a sufficient justification. The second is rooted in autonomy: Smokers, and prospective smokers, do not sufficiently appreciate the health risks of smoking, and graphic warnings can promote a kind of "debiasing" that statistical information fails to provide.<sup>127</sup> To this point, it might be added that government regulation is hardly being imposed on a blank slate. Recall that efforts to promote smoking involve a high degree of manipulation -- portraying happy, attractive smokers -- and the government can legitimately respond. In light of the number of lives at risk and the underlying evidence, these kinds of justifications do seem sufficient in the particular context of smoking.

## **B. Regulating Manipulation**

Should government regulate manipulation? In the context of political speech? Commercial advertising? We could imagine, or find, egregious cases in which it is tempting to say that it should. But the first amendment barriers are severe.

1. *Political speech and public figures.* In the context of political speech, the leading case is *Hustler Magazine v. Falwell*,<sup>128</sup> where the Court said that the first amendment protects a parody, depicting Protestant minister Jerry Falwell as engaged in an incestuous act with his mother at an outhouse. The parody was satirical; it can also be seen as a form of manipulation, designed to lead readers to see Falwell as a ridiculous figure and also a hypocrite. In the terms used here, the parody was an effort to appeal directly to System 1, so that people would not be able to regard Falwell in the same light in the future.

---

<sup>126</sup> See Jolls, *supra* note.

<sup>127</sup> *Id.*

<sup>128</sup> 485 US 46 (1988).

The Court unanimously ruled that the first amendment protected the parody. The Court acknowledged that to prevent genuine harm, states could regulate false statements of fact, which “are particularly valueless; they interfere with the truth-seeking function of the marketplace of ideas, and they cause damage to an individual's reputation that cannot easily be repaired by counterspeech, however persuasive or effective.”<sup>129</sup> But satire must be treated differently. “Were we to hold otherwise, there can be little doubt that political cartoonists and satirists would be subjected to damages awards without any showing that their work falsely defamed its subject.”<sup>130</sup> Even the most outrageous forms of satire are protected, because the idea of outrageousness has “an inherent subjectiveness about it which would allow a jury to impose liability on the basis of the jurors' tastes or views, or perhaps on the basis of their dislike of a particular expression.”<sup>131</sup>

To be sure, the Court’s reasoning was not unbounded. It acknowledged that the “actual malice” standard – allowing recovery of damages for statements known to be false or made with reckless indifference to the question of truth or falsity<sup>132</sup> -- would apply if readers had taken the parody as depicting actual facts.<sup>133</sup> But readers could not so take the parody here. This point leaves open the possibility that even in the political domain, certain forms of manipulation could be regulated if readers or viewers were affirmatively misled. But in view of the fact that the Court has pointedly declined to create a *general* exception to the first amendment even for false statements of fact,<sup>134</sup> any effort to regulate manipulative speech in the political context would run into severe trouble.<sup>135</sup>

With the design of any restrictions on such speech, there are independent questions of vagueness and overbreadth. If a government wants to prohibit the most egregious forms of manipulation in the political context, what, exactly, would it say? I have ventured a definition of manipulation here, but it is not exactly easy to adapt that definition to fit a provision of civil or criminal law. Manipulation has too many shades – which means that any effort to restrict it would likely be too vague and too broad.

---

<sup>129</sup> Id. at 52.

<sup>130</sup> Id. at 53.

<sup>131</sup> Id. at 54.

<sup>132</sup> *New York Times v. Sullivan*, 376 U.S. 254 (1964).

<sup>133</sup> 485 US at 56.

<sup>134</sup> *United States v. Alvarez*, 567 US – (2012), emphasizing that “some false statements are inevitable if there is to be an open and vigorous expression of views in public and private conversation.”

<sup>135</sup> At a minimum, it would be necessary to show that the manipulative statement created serious harm, and in the political context, such a showing would be highly unlikely to be enough, given the general commitment to the principle that the best correction, for arguably harmful speech, is more speech rather than enforced silence. See *Whitney v. California*, 274 U.S. 357, 374 (1927) (Brandeis, J., dissenting).

2. *Commercial speech*. The context of commercial advertising is different, because the burden on regulators is lighter.<sup>136</sup> But here as well, the first amendment obstacles are formidable. So long as the relevant speech is neither false nor deceptive, the government would need a powerful justification for imposing regulation.<sup>137</sup> The definitional issues remain severe, and even if they could be resolved, it would be plausible to say, in the general spirit of *Hustler Magazine*, that the marketplace of ideas is full of efforts to appeal to System 1, and to downplay or bypass deliberation and reflection.

Even if the commercial sphere is less immune from speech regulation, it is emphatically a place where manipulation is pervasive. The hope is that consumers will understand that advertisements are generally self-serving and that the process of competition will provide a sufficient corrective. To be sure, behavioral economics has raised serious questions about the realism of that hope.<sup>138</sup> But it is highly doubtful that those questions provide a sufficient basis for a general “manipulation exception” to the existing protection accorded to commercial speech.

### C. Consumer Protection

None of these conclusions mean that narrower forms of regulation could not be imagined. In the context of consumer financial products, various forms of manipulation are a widespread problem. Indeed, manipulation can be seen as a defining motivation for recent regulatory initiatives.<sup>139</sup> The Dodd-Frank Wall Street Reform and Consumer Protection Act states that the Consumer Financial Protection Bureau (CFPB) should ensure that “markets for consumer financial products and services are fair, transparent, and competitive.”<sup>140</sup> It calls for attention not only to “unfair and deceptive” acts and practices but also to “abusive” ones,<sup>141</sup> which can be seen as a reference to the worst forms of manipulation. In monitoring the relevant markets, the CFPB must consider the “understanding by consumers of the risks of a type of consumer financial product or service”<sup>142</sup> – a phrase that can easily be taken to reflect a concern about manipulation.

Implementing these requirements, the CFPB has adopted as its slogan, “know before you owe,” and its various efforts to ensure informed choices can be understood as an attack on manipulation as I have understood it here.<sup>143</sup> In consumer markets, of course,

---

<sup>136</sup> See note supra. For a valuable discussion from an economic perspective, see Richard Craswell, *Regulating Deceptive Advertising: The Role of Cost-Benefit Analysis*, 64 *Southern Cal. L. Rev.* 549 (1991).

<sup>137</sup> *Central Hudson Gas & Electric Corp. v. Public Service Commission*, 447 US 557 (1980).

<sup>138</sup> See Oren Bar-Gill, *Seduction By Contract* (2011).

<sup>139</sup> See Bar-Gill, *supra* note; Bubb and Pildes, *supra* note.

<sup>140</sup> 12 USC 5511

<sup>141</sup> *Id.*

<sup>142</sup> 12 USC 5512.

<sup>143</sup> See <http://www.consumerfinance.gov/credit-cards/knowbeforeyouowe/>

one problem is complexity, which can defeat understanding.<sup>144</sup> But another problem falls in the general category of manipulation, as in the form of “teaser rates” and various inducements that fall short of deceit, but that emphatically prey on System 1.<sup>145</sup>

A short, simple credit card agreement, of the sort provided by the CFPB, can be seen as a direct response to the risk of manipulation<sup>146</sup> -- and as an effort to ensure that System 2 is firmly in charge. Proposals to ban or restrict teaser rates can be understood in similar terms.<sup>147</sup> In cases of this kind, there is ample room for considering the problem of manipulation in deciding how best to regulate financial products. It is important to see that in such cases, government is regulating commercial practices, not advertising, and that its real concern is with practices that do not sufficiently trigger reflective deliberation on the part of consumers. We have seen that this is far from a self-defining category, but the CFPB’s initiatives can be taken as initial efforts to specify it.

### **Conclusion**

A statement or action can be counted manipulative to the extent that it does not sufficiently engage or appeal to people’s capacity for reflective and deliberative choice. Some forms of manipulation are egregious, as where a vivid, graphic description of an outcome (winning the lottery, dying in an airplane crash, losing a child) is invoked in order to convince people to engage in certain conduct (to buy a lottery ticket, to take a train, to buy extra life insurance). Some arguable forms of manipulation are mild, as when a politician, an employer, or a waiter uses loss aversion, tone of voice, and facial expressions to encourage certain decisions. Thus defined, manipulation is a pervasive feature of human life. It is for this reason that while the legal system is generally able to handle lies and deception, it has a much harder time in targeting manipulation.

In their troublesome forms, manipulative acts fail to respect choosers; they undermine people’s autonomy and do not respect their dignity. The welfarist objection, rooted in the idea that choosers know what is in their best interests, is that when they are products of manipulation, people’s choices may not promote their own welfare, precisely because choosers have not been put in a position to deliberate about relevant variables and values. This is likely to be true if the manipulator is ill-motivated, but it might also be true because the manipulator lacks relevant information.

From the welfarist point of view, manipulation is only presumptively disfavored. A benign, knowledgeable manipulator could make people’s lives go better and possibly much better. But under realistic assumptions, the presumption against manipulation is

---

<sup>144</sup> See Bar-Gill, *supra* note.

<sup>145</sup> See Ryan Bubb and Richard H. Pildes, *How Behavioral Economics Trims Its Sails and Why*, 127 *Harv. L. Rev.* 1595, 1661-1662 (2014).

<sup>146</sup> <http://www.consumerfinance.gov/credit-cards/knowbeforeyouowe/>

<sup>147</sup> Oren Bar-Gill and Ryan Bubb, *Credit Card Pricing: The Card Act and Beyond*, 97 *Cornell L. rev.* 967 (2012).



justifiably strong, because manipulators are unlikely to be either benign or knowledgeable.