



A RESTRICTION MAYBE, BUT IS IT  
PATERNALISM?  
COGNITIVE BIAS AND CHOOSING  
GOVERNMENTAL DECISION AIDS

J.D. Trout\*

**1. Introduction**

In “Paternalism and Cognitive Bias,”<sup>1</sup> I survey cognitive biases, and conclude that they routinely prevent us from exercising our autonomy. I argue that these obstructions can be sufficiently serious to warrant institutional intervention. Finally, I argue that in many of these cases, the institutional intervention is not paternalistic.

In this essay I address three issues that have arisen in critical response to that paper. The first concerns how to define the proper scope of interventions deemed paternalistic. The government routinely restricts our behavior by promising to sanction us if we engage in certain actions. Only the most extreme libertarian<sup>2</sup> — and I hesitate even to use that dignified label for some of these ob-

---

\* Professor of Philosophy and Adjunct Professor at the Parmlly Hearing Institute, Loyola University Chicago.

<sup>1</sup> J.D. Trout, *Paternalism and Cognitive Bias*, 24 LAW & PHIL. 393 (2005).

<sup>2</sup> I have in mind the kind of “militia” member described in STEPHEN HOLMES & CASS SUNSTEIN, *THE COST OF RIGHTS: WHY LIBERTY DEPENDS ON TAXES* (1999).

jections—would regard all such constraints as paternalistic. Second, many restrictions that at first appear “paternalistic,” in that they limit a person’s range of choices “for his own good,” may not, in fact, do so “against his will”—thus raising the question of whether such restrictions are really paternalistic at all. Third, I will examine whether long-term goals should be given priority over short-term goals. I will close by considering the value of being explicit about our social priorities.

## 2. The Meaning of “Paternalism”

Some might find puzzling the idea that institutional intervention in decision-making is not necessarily paternalistic. To these people, institutional intervention on behalf of a person, putatively for that person’s own good, defines paternalism (whereas others might simply call it good government). But to philosophers, paternalism means something quite specific, i.e., “the interference of a state or an individual with another person, *against their will*, and justified by a claim that the person interfered with will be better off or protected from harm.”<sup>3</sup>

There are many other definitions of paternalism, but most are captured by this core account. “Against one’s will,” can be interpreted narrowly or broadly. On the narrow interpretation, “against one’s will,” simply means an interference with one’s current desires; so as long as the other conditions are in place, an intervention is paternalistic if the government interferes with any old preference you are currently in the grips of. On the broad interpretation, an intervention is against one’s will only if it interferes with all of the desires one would avow if asked.

One kind of paternalism, so-called “soft” paternalism, applies to protected groups, such as children or adults who are mentally incompetent.

---

<sup>3</sup> Gerald Dworkin, *Paternalism*, in THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2002) (emphasis added), available at <http://plato.stanford.edu/archives/win2002/entries/paternalism/>.

Soft paternalism holds that the state has the right to prevent self-regarding harmful conduct (so far as it *looks* paternalistic) *when but only when* that conduct is substantially nonvoluntary, or when temporary intervention is necessary to establish whether it is voluntary or not.<sup>4</sup>

Is government intervention justified in the case of the decision-making conduct of these two groups (children and mentally incompetent adults)? Feinberg offers a test based on the kinds of conditions that make choices substantially nonvoluntary, including “[I]gnorance, coercion, derangement, or other voluntariness-vitiating factors.”<sup>5</sup> Such a test is designed to protect the cherished values of voluntariness, free will, and autonomy. If vitiating factors predominate, then the person’s conduct “does not come from his own will, and might be as alien to him as the choices of someone else,”<sup>6</sup> then, intervention may be justified.

But voluntariness is characterized by degrees, not bright lines. This can be troublesome in certain cases—for example, we don’t allow sixteen year-olds to vote, even though some of them may be more competent to process the relevant political information than many who can legally vote. The need for clarity in administering the law calls for a bright-line rule; thus, we deem minors unqualified to vote, regardless of their actual qualifications. The same reasoning may apply if the cognitive biases of ordinary, mentally competent adults can be considered “voluntariness-vitiating”—that is, if, under normal conditions, these biases can and do prevent us from fully and clearly deliberating about some goal (compared to, say, the deliberative process we would follow were we free of bias). Although this is not an easy question to answer, many decisions affected by cognitive biases result in behavior that would appear “not entirely voluntary,” or simply “nonvoluntary.”

---

<sup>4</sup> JOEL FEINBERG, HARM TO SELF 12 (1986).

<sup>5</sup> *Id.*

<sup>6</sup> *Id.*

If this is the case, correction of cognitive bias may not be paternalistic, even when the cognizer is otherwise competent. An intervention is not against one's will simply because one wishes at a given moment that he were not bound by the intervention. Goals are complex enough that I may simultaneously desire at least two goals that conflict. I may, for example, both want to retire with more money rather than less, and eat a pastry every morning. If so, the practice of buying a pastry every day instead of investing that money for retirement is inconsistent *with my expressed goal* of saving money. It is quite clear that I have both desires. I really, really want the pastry, and I really want to retire with more money rather than less.

People also have second-order desires, which is why the ambiguity of the "against one's will" clause in standard forms of paternalism becomes important. People who have the first-order desire to retire better rather than worse off also have a second-order desire that they have the capacities of judgment and willpower—whatever they are—to make that happen. And the fact that, in many situations, our reasoning capacities underperform government prosthetics creates a conflict between first and second-order desires.

There are many political postures that are deeply conflicted about life in civil society. Social and political philosophers have long recognized the difference between not liking the government's role in a process, and the more specific charge that its role is paternalistic. For the same reason, they have recognized that a competent person cannot be the subject of paternalism if they agree with the measure they are forced to abide by. Is a helmet law paternalistic if everyone subject to it agrees they should wear the helmet? To answer, we need to distinguish between a distaste for paternalism proper, and the more general distaste for governmental constraints on conduct. If you oppose this measure, yet you clearly consent to its intended outcome, it is because you don't like governmental regulation, not because that governmental regulation is paternalistic.

In addition, as many scholars have noted, there is probably little human behavior to which Mill's principle of liberty properly applies. Genuinely self-regarding behavior is hard to come by, and it is not saying much to say that the helmet law, for example, is consistent with Mill's principle, properly applied. Getting killed on a motorcycle at significantly higher rates concerns not just the rider but the rider's family and friends. Further, it affects not just those in the rider's social orbit, but the stranger under whose car he slid, and who now must live with the fact of the rider's death, and so on. This example is not designed to show that the federal government can legitimately legislate that all motorcyclists wear helmets; rather, it is designed to show that, if the government *can't*, it is not because Mill's principle of noninterference compels its abstention. On the contrary, those beyond the boundary of purely private interest are influenced in ways that Mill failed to explore, a point conceded by friend and foe of Mill's social philosophy.

Significantly for the present discussion, Mill wrote that the non-interference principle applied to all but two groups of people: children, and people living in "backward states of society". These two exceptions involve compulsions that make an individual's behavior effectively involuntary. Such is also the case with cognitive biases. It may be wise, then, to identify the frailties in judgment associated with these biases and take steps to protect people from them. In each of these cases—because behavior cannot be fully informed by independent reason—it might be rational and desirable for people to yield to governmental assistance in decision-making.

Life in a complex, civil society is not chiefly about accommodating the strongly-felt desires of clinically alienated individuals at the expense of the majority. Instead, it is about affording, and where necessary, administering the common good. True, we can always create a character who is a chronically miserable social isolate with a death wish—one who sees in a helmet law the end of freedom as we know it. It is not surprising that certain isolated groups are not burdened with the empathic insight that their choices, at first apparently self-regarding, can actually harm others. This inability to empathize is common among sociopaths, but ordi-

nary people suffer failures of empathy more often than we suppose.<sup>7</sup>

### 3. Paternalism and the Contracts of Consenting Agents

The standard conception of a contract is an agreement between two or more parties, especially an agreement that is written and enforceable by law. It is a time-honored principle in the U.S. that, to the extent possible, government institutions should not intervene to prevent "consenting agents" from entering into bargains with one another of their own "free will". There are raging controversies about the nature and limits of both consent and free will. If we could identify those limits, we might be better able to determine what the "extent possible" is.

Even the smartest consumer will likely lack the arcane knowledge of biochemistry to make drug-prescribing decisions, or arcane knowledge of digestion to set food contamination limits. That is why we have a Food and Drug Administration. So it is not a question of whether our decisions *should be* regulated; they are regulated already.

Included in the idea of a contract is that, to the extent possible, a free market should determine our opportunities, entrepreneurial or otherwise. If a product's name makes it more recognizable, then, to the extent possible, the naming decision should not be fettered by non-market forces. To take a recent example, the perception is that drugs that begin with 'z' are more technical-sounding, and so more desirable to the public. Drug companies love them, and customers may have more favorable attitudes toward them. Among drug companies then, 'z' names are much admired. So is not killing patients. But doctors, nurses and pharmacists are much more likely to commit a prescription error when the word-initial sequence (in this case, of a drug name) is in lexically crowded territory. Now there is a Council on Drug Naming that the FDA uses to

---

<sup>7</sup>J.D. TROUT, (Forthcoming) (unpublished book manuscript, Viking/Penguin).

name new drugs, so that they are less likely to be confused with existing names.

Consider also the regulatory action of the FDA. The FDA routinely sets limits on the acceptable “adulteration” of food products; I have selected peanut butter as my example, but could have selected many other delectables regulated by the FDA. The term “adulteration” is a surprisingly neutral term for insect parts, rodent hairs, rodent feces, and maggots, as you will find in FDA regulations. The regulation for peanut butter “Sec. 570.300 - Peanut Butter - Adulteration with Filth; Grit (CPG 7112.03)” says that the following conditions, among others, warrant seizure or citation by the Division of Compliance Management and Operations:

#### Filth

The peanut butter contains an average of 30 or more insect fragments per 100 grams; or

The peanut butter contains an average of 1 or more rodent hairs per 100 grams.<sup>8</sup>

Is it a paternalistic provision that would cite or seize shipments at or above these proscribed levels of adulteration? Shouldn't I be free to enter into agreements, to make contracts that would exploit a niche in the market for cheaper peanut butter? And shouldn't a public, hungry to trade a few insect fragments for a cheaper peanut butter, be permitted to purchase what I have to sell?

If you are inclined to answer that the FDA's proscription is paternalistic, then it is probably time to ask: Are governmental regulations of this sort *ever* warranted? How about non-contract issues (at least for the consumer), like seat belt laws? How about the contract issue of the sale of toxic chemicals? Is it legitimate for the

---

<sup>8</sup> Office of Regulatory Affairs, U.S. Food and Drug Admin., FDA/ORA Compliance Policy Guides § 570.300 (2006), available at: [http://www.fda.gov/ora/compliance\\_ref/cpg/cpgfod/cpg570-300.html](http://www.fda.gov/ora/compliance_ref/cpg/cpgfod/cpg570-300.html) (last visited Sept. 9, 2007).

government to interfere with a contract in the interest of, say, equal opportunity or anti-discrimination?

The obvious move for the anti-paternalist would invoke Mill's principle of interference. But once again, the upshot of social and political scholarship is that it is difficult to find purely self-regarding behavior; it is difficult to tell whether the effects of an individual's conduct do or do not affect others in foreseeable ways. Purely self-regarding behavior, then, can be treated as a methodological fiction. Hence, as a procedural matter, it would be easier to address such issues if all parties were first clear about whether *any* regulation is *ever* legitimate, and if so, under what circumstances. Silence, unclarity, coyness, or deception about one's initial commitments turns every discussion of government regulation into a referendum on whether the institutions with which we have socially contracted—offices of the government—should ever have any power to constrain our conduct.

If the impermissibility of governmental constraint is supposed to issue from certain basic rights, then at least we can highlight the fact that these basic rights have a mobile and amorphous history. When we attempt to determine the scope of legal contract, we can choose to acknowledge or ignore the standard qualifications that make it enforceable by law—that one or both of the parties be of legal age, and not be insane or a drunkard—i.e., they must be competent. We can also choose to acknowledge or ignore that qualifications like the competence constraint have changed over time. Under Roman Law, for instance, deaf people couldn't witness wills. In the antebellum South, Blacks couldn't enter into contracts, as Chapter VII of the American Slave code makes clear.<sup>9</sup> Did lawmakers simply consult their intuitions when devising these laws? If so, isn't the decision-making process still roughly the same

---

<sup>9</sup> WILLIAM GOODELL, *THE AMERICAN SLAVE CODE IN THEORY AND PRACTICE: ITS DISTINCTIVE FEATURES SHOWN BY ITS STATUTES, JUDICIAL DECISIONS, AND ILLUSTRATIVE FACTS* (1853), available at <http://www.dinsdoc.com/goodell-1-0a.htm> (last visited March 10, 2007) ("Being held as Property, and incapable of making any Contract, [slaves] cannot contract Marriage recognized by Law.").

today? Isn't it true that lawmakers still rely on their intuitions and what they believe is casually inspectable evidence when determining what constitutes competence? If it is, then what makes today's contract laws any less datable if they fail to account for cognitive biases?

As is apparent from the above, we can define freedom of contract as liberty among consenting adults to agree to terms of exchange, and then choose to define "consenting adults" in a way that suits our tastes. Current tastes among "contract libertarians" ignore findings about (involuntary) cognitive bias. But ignoring important evidence and linguistically legislating the result seldom sheds light. Opponents of governmental correctives for cognitive bias should address the most obvious costs of cognitive bias. They might begin with this sequence of seven questions:

- (1) Are we as competent as anyone to judge technical material, such as literature on discounting the future, or to identify structures in the mental lexicon that might cause prescription errors?
- (2) If not, is the private hire of experts practical and available to all?
- (3) Is evidence/information for our decisions typically freely available?
- (4) Are the information search problems facing a rational person computationally tractable?
- (5) Are we capable of placing an informed bet that we are competent to make a specific judgment (given our lack of arcane knowledge, and our general cognitive bias of overconfidence)?
- (6) Are we ever justified in imposing an institutional constraint for the purposes of coordinating activity (for example, to avoid problems that would arise if doctors, as experts, were exempt from certain FDA regulations in their personal lives, allowing them but not others to open a business selling peanut butter with a higher proportion of insect parts, or import it from other coun-

tries? Or, to use another example, could smart 16 year olds be allowed to vote)?

- (7) Does the very existence of the FDA, FCC, etc. indicate our recognition that institutions are needed to overcome information-processing problems that arise from cognitive boundedness?

#### 4. The Intervention Test, and the Confusion over Long- and Short-Term Goals

Were we to treat all instances of government regulation of self-regarding behavior necessarily as instances of paternalism, we would risk trivializing the notion of paternalism altogether, stripping it of any meaning independent of regulation and intervention. But once properly separated, we can first ask whether a particular intervention is paternalistic, and then whether the intervention is justified.

Our government requires us to contribute to Social Security, a form of pension system. The government prohibits the sale of assorted drugs considered harmful, and regulates the names of prescription drugs on the market. The government forbids consent to certain forms of assault to be a defense against prosecution. Some of our state governments legislate that motorcyclists must wear helmets. The list of government-regulated actions could be continued at length.

Clearly, the notion of self-government is central to our conception of autonomy.<sup>10</sup> Due to this centrality, there should be a strong presumption against interfering with individuals' considered judgments about their best interests, about what decisions will be most conducive to their well-being. But we still need to know when the government can legitimately regulate our conduct. In "Pa-

---

<sup>10</sup> FEINBERG, *supra* note 4, at 28. Feinberg's book contains perhaps the most influential modern discussion of paternalism. A very useful discussion of paternalism and its legal implications can be found in LAWRENCE O. GOSTIN, *PUBLIC HEALTH LAW: POWER, DUTY, RESTRAINT* (2000).

ternalism and Cognitive Bias” I introduced a test for warranted intervention:

The intervention is warranted if, against a background of a fully informed decision-maker and an unbiased standard, the decision-maker would not have made the decision he or she did. In the present case, if the decision-maker knows that they can’t resist taking an unreasonable risk, they would have consented to any number of low-cost corrective measures.<sup>11</sup>

When a regulation assists the person in effectuating his or her long-term goals or plans, there is no relevant interference, and the regulation is not paternalistic; that is, the measure is not, in any relevant sense, “against the will” of the agent.

Like most general standards, this one deserves some fleshing out. What is meant by “a fully-informed decision-maker”? One thing I do not mean is a decision-maker who implements a distinctly counterfactual normative ideal, like LaPlace’s demon or an optimizing Bayesian model. If the standard is to be effective, it may not be able to simply take people as they are—cognitive warts and all—but it should take them as they can be in a reasonable amount of time, with a reasonable amount of effort, with a normal standard of intelligence.

Thus, a fully-informed decision-maker is one who has all of the available information relevant to their decision, *including information about the ways in which they are prone to bias*. For example, a fully informed decision-maker is one who, among other things, knows, as the research shows, that they systematically err in decisions about what will make them happy. Accordingly, they will, in general, discount the expectation that a new material acquisition of a new material good, like a car or a coffee maker, will durably contribute to their happiness.

---

<sup>11</sup> Trout, *supra* note 1, at 413-14.

Legal scholars have, like epistemologists, become accustomed to evaluating the quality of a position by examining particular decisions. In our recent book, *Epistemology and the Psychology of Human Judgment*,<sup>12</sup> Michael Bishop and I argue that individual decisions are always subject to natural variation, so it is hard to know what to conclude when a particular token belief proves unjustified. Instead, we should seek epistemic excellence, which consists of the efficient allocation of cognitive resources to robustly reliable reasoning strategies, once applied to significant problems.

The relation between liberty, choice, and welfare is a complicated one. But the balance of legal doctrine and political philosophy deny that unfettered liberty and unlimited choice reign absolute. Neither liberty nor choice is an unconditional value. In "Paternalism and Cognitive Bias" I argued that cognitive biases are common, spontaneous, and costly. Because of their pervasiveness and intractability, the damage they cause is bound to be great, and costly (at the start) to reverse. And like cognitive practices with costly and potentially self-destructive effects, the costs may be great enough to warrant regulating an individual's choice about a largely self-regarding issue.

Most of our biases are automatic, and are very vulnerable to the on-the-spot effects of hot cognition. Our discounting biases are too.<sup>13</sup> Correctives to these most common and damaging biases may seem to privilege long-term over short-term goals. They counsel that we slow down and consider hidden costs, place distance between a desired object and our behavior, and order our overarching priorities. This may seem to defer immediate gratification, but this is a side effect. What these strategies are really privileging is the correction of more serious rather than less serious biases.

It is uncommon for people to oversave, so we seldom have to privilege short- over long-term desires. As a contingent, empiri-

---

<sup>12</sup> MICHAEL BISHOP & J.D. TROUT, *EPISTEMOLOGY AND THE PSYCHOLOGY OF HUMAN JUDGMENT* (2005).

<sup>13</sup> J.D. Trout, *The Psychology of Discounting: A Policy of Balancing Biases*, 21(2) Pub. Aff. Q. 201 (2007).

cal matter of fact, our short-term desires appear to swamp our ability to defer gratification. So we may need to render ineffective those mechanisms that disable our resistance to the tempting morsels of life that intoxicate us at the moment, but not because our long-term goals always represent the voice of reason, or that we always accurately identify what is best for us in the long run. Instead, sometimes long-term goals should prevail, sometimes short-term ones. And incentives should be structured appropriately.

My arguments appear to privilege long-term over short-term goals for two contingent reasons: 1) When people are called upon to conceptualize the things that matter to them in the long run, they do so in more general terms than their avowals about short-term goals, in which whims and visceral desires dominate, and 2) When people are called upon to conceptualize their long-term goals in specific terms, they get things very wrong. We are badly prone to the durability bias

General characterizations of our goals allow us to see how humans are similar. When asked specific questions about skills and goals, people tend to focus on the attributes and pursuits that separate themselves from others. We find this in research on pluralistic ignorance, on risk, and in affective forecasting.<sup>14</sup> Especially in the case of affective forecasting, people badly mis-predict the events that will matter to their long-term well-being. Subjective well-being research shows that people mistakenly believe that their happiness depends on things like getting that vacation home in the Berkshires, a new sports car, or a higher salary. But these changes are short-lived, and we invariably adapt to their impact.

Notice what else happens when we ask more general questions. Many people wouldn't make sacrifices to get a sports car, or trade peace of mind for a new mortgage on the vacation home. But everyone would rather retire with significantly more than less money, and live a longer rather than shorter life (all other things

---

<sup>14</sup> DANIEL GILBERT, *STUMBLING ONTO HAPPINESS* (2006); *see also* Christopher K. Hsee & Elke U. Weber, *A Fundamental Prediction Error: Self-other Discrepancies in Risk Preference*, 126 J. OF EXPERIMENTAL PSYCHOL.: GENERAL 45 (1997).

being equal)—and many would make sacrifices to meet these goals. When it comes to the general things we want in our lives, we aren't really that special. This is natural, because our ignorance of the remote future is nearly complete. So the best that most people can shoot for is more rather than less retirement money, etc. But people will focus on vastly *different* activities in the next 20 days.

So my account of debiasing doesn't actually privilege long-term desires. Instead, it suggests the importance of ordering our goals and in doing so, directs our attention to long-term influences on our ultimate well-being—precisely because they have been neglected. It is not presumptuous to suppose that a person wants to retire with more rather than less money, and so to structure incentives or options to encourage that end. But it is presumptuous to suppose that in the next 20 days someone can decide exactly what location they will want when retired, and which home they will want to live in. We are not very good at forecasting our affective reactions to specific future events. The evidence of affective forecasting errors radically changes the structure of this dispute, because the second condition required for paternalism—that a regulation be “against one's will”—is now virtually uninterpretable. The response to it is now best understood as expressing the much weaker view—and also much darker and mysterious—that we are entitled to our own mistakes.

Consider one last intervention, in the form of building code or zoning regulations designed to yield crime-resistant dwellings.<sup>15</sup> There are objections to situational crime control, architectures that discourage vandalism and assault. One objection is that it delivers individuals from the responsibility for their choices, producing a kind of cognitive hazard for the potential victim. Another is that the control measures are insidious, because imperceptible. In the modern world, we must choose the intrusions that are least offensive. Is it better to be protected from crime by managing architecture, or be hypnotized into buying a product by a store environment that ma-

---

<sup>15</sup> See, SITUATIONAL CRIME PREVENTION: SUCCESSFUL CASE STUDIES (Ronald V. Clarke ed., 1992).

nipulates your suggestibility by blink rate? Unlike the sneaky advertising gimmicks designed to disable our considered judgment and get us to buy products we don't want, passive schemes for crime prevention can do some good.

## **5. Conclusion**

I have made three arguments. First, I have argued that the scope of paternalism is unclear, but governmental restrictions on conduct introduced for our own good are routine. Second, I argued that, due to conflicting desires, such restrictions are not necessarily "against one's will," despite spontaneous avowals to the contrary. Hence, it is unclear whether such intervention is really paternalistic at all. Third, I argued that it is not long-term goals that should be given priority over short-term ones, but important values over less important ones. So there is great value in being explicit about our priorities, especially where strategies that correct the biases may interfere with our autonomous pursuits.