

PATERNALISMS AND NUDGES

DANNY SCOCCIA*

Abstract: Richard Thaler and Cass Sunstein (T/S) have defended ‘nudges’ aimed at smoking, overeating, etc. as a ‘means’ paternalism that leaves its targets ‘better off as judged by themselves’. Their libertarian critics have charged that these behaviours are often perfectly rational and that the nudges would ‘impose’ on their targets ‘ends’ that they reject. This paper argues that whether or not the behaviours are rational is difficult to say, but the critics are right in claiming that T/S fail to take seriously their targets’ true preferences. This is evident, in particular, in Sunstein’s recent reply to an ‘autonomy objection’ to nudging. The upshot is that the nudging paternalism T/S support cannot be defended as a means paternalism that is deferential to its targets’ own ends or values. The only way to defend it is via a ‘prudential’ paternalism that, given the preferences of many of its targets, will often be ‘ends’ paternalism.

Keywords: paternalism, rationality, prudence, nudge, autonomy

Paternalistic actions, laws, or policies are often thought to use coercive means to interfere with the choices of other persons for what the paternalist believes is their own good. Richard Thaler and Cass Sunstein (hereafter, T/S) have defended a ‘libertarian paternalism’ that uses different means to achieve a different end (Thaler and Sunstein 2003, 2008). The means it uses, ‘nudges’, are supposed to be either minimally coercive or not coercive at all. Moreover, its aim is not to leave its targets better off by *the paternalist’s* own lights, but instead to leave them better off ‘as judged by themselves’.

What is a ‘nudge’? T/S define it as a ‘factor that alters the behavior of Humans, even though it would be ignored by Econs’ (Thaler and Sunstein 2008: 8). The choices of an Econ (an idealized person stipulated to possess full epistemic and economic rationality and to be fully informed about her options) are determined by the ‘content’ of her option set

* 2734 S. Las Palmas St., Mesa Arizona 85202, USA. E-mail: dscoccia@nmsu.edu. URL: <https://philos.nmsu.edu/faculty-and-staff/danny-scoccias-home-page/>

together with a utility function that obeys the axioms of rational choice theory. This means that nudges have to be substantially non-coercive, because coercion increases the costs attached to some of one's options and Econs are responsive to changes in costs.¹ Nudges can influence the choices of ordinary Humans only because they, unlike Econs, have certain propensities to irrationality hardwired into their psychology. Examples of nudges include mandatory waiting or 'cooling off' periods to implement certain types of choice to reduce the chances that one will choose irrationally due to excitement or impetuosity, and making employee contributions to a retirement savings account the default option (the one that one has to opt out of if one doesn't prefer it), which increases saving by exploiting status quo bias or inertia. Insofar as a warning discloses information about costs/benefits that one didn't have, it is not a nudge, though the way it is framed (e.g. use of skull and crossbones symbol rather than plain text to warn that a substance is poisonous) might involve one.²

T/S's paternalistic 'choice architect' decides which nudges to use by determining which will leave their targets better off 'as judged by themselves'. What does that mean? Many of our choices/'revealed preferences' are context-dependent, endogenous, and thus malleable – change the choice environment with nudges here or there and that will predictably elicit different choices/'revealed preferences'. If all of our preferences were like that, T/S's choice architect would have no criterion for determining which nudges leave their targets better off 'as judged by themselves'. Thus, T/S must assume:

- (1) 'Underneath' our endogenous, context-dependent preferences or choices are some deeper context-independent preferences.
- (2) One's 'welfare' is properly determined by the extent to which those 'true' preferences (or some subset of them) are satisfied, not by the paternalist's preferences or opinions about one's good.
- (3) Because we are ordinary Humans rather than Econs, we are often *irrational* and fail to make the choices best calculated to satisfy those preferences.

¹ Sunstein has noted, correctly I think, that the distinction between a nudge and coercive shove is not sharp but rather one of degree. This is especially true if, as he supposes, some ways of 'framing' information impose 'psychic' (rather than 'material') costs on those receiving it.

² T/S sometimes say that warnings about the risks of smoking, not wearing a seatbelt, etc. are nudges, but this is at best misleading. Even if one assumes that Econs are already in possession of all such information (and one shouldn't assume that for a reason to be noted later), that only makes warnings addressed to Econs unnecessary; it doesn't mean that the information itself would have no influence on their choices.

Their defence of the paternalistic state nudges that they favour also assumes:

- (4) The 'deeper' preferences referred to in #1 include good health, a normal lifespan, and a comfortable retirement. The choices we would make if we were Econs to better satisfy those preferences include trying to quit smoking (if we're smokers), reducing our intake of high sugar, high trans-fat foods, saving more for retirement, etc. Moreover, the experts and officials in the State Office for Nudging can and do know that these are the choices it is rational for us to make given our deeper preferences.
- (5) There is nothing *intrinsically* or deontologically wrong with the state's nudging its citizens; its acceptability is determined by its costs and benefits.
- (6) The benefit to cost ratio of state nudging will be positive so long as it is effective at inducing its targets to make more healthful and financially more prudent choices. That is, it carries no other significant costs or risks that might outweigh that benefit.

Unsurprisingly, there are critics of each of these assumptions.³

My discussion in what follows is divided into five sections. In the first I spell out the objection to #4 more carefully, and I concede to its defenders that many of the choices that T/S's nudges would target might be rational. In [Section 2](#) I distinguish four kinds of paternalism (means/prudential, ends/prudential, means/non-prudential, and ends/non-prudential) on the basis of another, more basic distinction between one's 'conception of the good' and one's prudential good. I also distinguish beneficence, the *prima facie* moral duty to promote the prudential good of others, from the Deference to Others' Conceptions of the Good (DOCG) principle, the *prima facie* duty not to interfere with others' choices in pursuit of their morally permissible conceptions of the good. The claim that smoking, undersaving, etc. are not irrational may weaken the case for

³ Re #1: Robert Sugden objects that there is no 'fact of the matter' about what choices we would make if we were Econs (and thus, what our 'true preferences' are), because the irrational choices that we as ordinary Humans make are due to essential features of our psychology (Sugden 2008).

Re #5: A deontological objection to paternalism that uses nudges is that it is 'manipulative' and manipulation is wrong. The proper, respectful way to induce others to make better choices is honest, transparent rational persuasion (more on this objection in [Section 4](#) below).

Re #6: One objection is that being nudged to make better choices deprives one of the opportunity to make and then learn from one's bad choices; that makes at least excessive nudging infantilizing (Bovens 2008). Another objection is that T/S downplay the danger of a 'slippery slope': paternalistic nudging by the state today is likely to lead to paternalistic coercive shoving by the state tomorrow (Rizzo and Whitman 2009).

a means/prudential paternalism that targets them, but not the case for an ends/prudential paternalism that does so. Section 3 argues that while T/S's paternalism is supposed to be the means/prudential type, it is neither. It isn't prudential paternalism, because its aim is not consistent with that of any normative theory of prudential value, even a preferentist one. It isn't means paternalism, because T/S's libertarian critics are correct in their complaint that T/S attribute to many of the targets of their nudges preferences that they simply don't have. Section 4 explores the 'autonomy objection' to T/S's nudges and distinguishes two versions of it – a non-Kantian one based on DOCG and a non-Kantian one not based on it. Sunstein has a reply to the non-Kantian objection, but it fails. The success of that objection is another reason why T/S's nudges cannot avoid being an ends paternalism that violates DOCG. In a brief concluding section, I describe how one might defend prudential paternalism while avoiding Sunstein's mistakes. It involves admitting that it will often be ends paternalism but urging that that is no real objection to it, because DOCG should be rejected. Whether abandoning DOCG is too steep a price to pay to defend prudential paternalism is a question I make no attempt to answer in this paper.

1. ARE SMOKERS, OVEREATERS, ETC. IRRATIONAL?

Assumption #4 – that most smokers, overeaters, under-savers etc. are making irrational choices contrary to their 'true preferences' for good health, longevity, a comfortable retirement, etc. – invites a couple of different objections. Both concede that there is a 'fact of the matter' about what people's 'true preferences' are. The first denies that any government official is in a position reliably to know them, because she will be subject to biases, ideological preconceptions, pressures from special interest groups, and the like. The second goes further, claiming that the beliefs that T/S's paternalistic nudge has about her target's 'true preferences' are in fact false. T/S assume that there is substantial uniformity in the trade-offs that people would accept, if they were fully rational and well informed about their options, between health/longevity or a financially secure retirement, on the one hand, and myriad other goods (e.g. carnal pleasure here and now, rewarding interpersonal relationships, intellectual stimulation, etc.), on the other. But – the objection alleges – the different trade-offs that different people in fact accept via their actual choices, as well as their avowals, what they express regret for, etc. make it overwhelmingly likely that such an assumption is false. Many of us would continue to make the 'bad' lifestyle choices that nudging and coercive paternalists would target, even if we were fully rational and well informed about risks. William Glod presses this objection to the claim that overeating is irrational, Daniel Shapiro to the claim that smoking is irrational, Gerry

Gaus to the claim that riding a motorcycle without a helmet is irrational, and Gary Becker and Kevin Murphy to the claim that recreational hard drug use is irrational.

Consider first Glod, who says:

I challenge the claim that many people are *prima facie* irrational to eat enough to make themselves obese and thus tend to have shorter life spans with attendant maladies. Some really enjoy eating and are willing to pay the price; others overeat as a response to stressors, which may be a rational response to a difficult situation if healthier and comparably desirable channels are not readily available. We lack the information to make the claim of irrationality without knowing if all else is equal when a person performs unhealthier actions rather than healthier ones. (Glod 2015: 614–15)

Glod is claiming that rationality should be the default assumption, the burden of proof falls on those who would deny it in any particular case, and T/S don't satisfy that burden in the case of the persons they would target with their 'eat healthier!' nudges.

Daniel Shapiro says something similar about smoking. He rejects Robert Goodin's contention that most smokers 'falsely and irrationally underestimate the risks of smoking' (Shapiro 1994: 188). Shapiro cites a study by the economist W. Kip Viscusi showing that smokers (as well as non-smokers) in fact overestimate the risks of dying from lung cancer. He argues that smokers continue to smoke in spite of the risks to health/longevity because smoking satisfies a variety of desires. It is the fact that smoking serves so many different ends, making it well integrated into the smoker's life, rather than the fact that it is an 'addiction' requiring great feats of willpower to overcome, that explains why half of adult smokers who try to quit fail (Shapiro 1994: 194). Most importantly, the Viscusi study also showed that smokers may be less risk averse than non-smokers. Insofar as the choice to continue smoking is the result of risk-seeking preferences, there is nothing irrational about it.

Consider next Gaus's claim that even though helmetless riding creates a severe risk of brain damage if the motorcyclist is in a traffic accident, many riders prefer to run that risk so that they may enjoy the sensation of riding with 'the hair in the wind'. To 'middle-class, middle-age people, who are risk-averse' that trade-off seems foolish, but it is not irrational for those riders simply because they are much less risk averse (Gaus 2003: 23).

Finally, the 'rational addiction theory' of Gary Becker and Kevin Murphy holds that 'addictions, even strong ones, are usually rational in the sense of involving forward-looking maximization with stable preferences' (Becker and Murphy 1988: 675). As one critic of the theory describes it, the theory ('absurdly') assumes that:

By smoking today you increase your smoking stock. Stable preferences are defined over such stocks in addition to the consumption goods ...

Consumption of heroin, tobacco, jogging, and television watching are all characterized by having such effects on future levels of various stocks. Individuals have precise, quantitative estimates both of these effects and of all future levels of 'exogenous' factors such as prices, norms, and public regulations. To decide whether or not to smoke tobacco, cigarettes, heroin, or crack, whether or not to turn on the television, go for a jog or eat a chocolate bar, individuals exploit these effects and design a detailed consumption plan for their future life exhausting all gains from all trade-offs across time and goods conditional on opportunities and exogenous factors; i.e. they map out – in precise detail – their optimal, lifetime plans of moment to moment consumption of these goods. (Rogeberg 2004: 271)

Another critic of the Becker/Murphy theory, Jon Elster, argues that addictions to heroin, nicotine, gambling, etc. impair one's ability to form epistemically rational beliefs about the effects of the activity on future utility (Elster 1999). Even if a choice to gamble or use heroin always reflects a judgement about what it is all things considered best to do (i.e. is never due to weakness of will), it will still be irrational if the judgement depends on epistemically unreasonable beliefs. Elster claims that addicts tend to have such beliefs regarding risks, ease of altering future behaviour, etc., owing both to errors in causal and inductive reasoning ('cold' irrationality) and to cognitive dissonance and wishful thinking ('hot' irrationality) (Elster 1999: 26). Against Becker's rational addiction theory, Elster objects:

Becker tacitly assumes that the addicts' awareness of future consequences is not impaired – simply that the consequences are given less weight in the utility function. But how can he tell the difference? To be sure, by stipulating that the drug works on the discounting and not on the beliefs, he can still claim that the addiction is rational ... This cannot, however, be a sufficient reason for making the stipulation. What we observe is simply that the addict is less swayed by future consequences than he was before he took up the drug. (Elster 1999: 44–5)

Is Elster right? Do drug use, smoking, etc. only affect one's utility function, in which case they are not irrational, or do they also directly affect one's beliefs (via cognitive dissonance, wishful thinking, and so on) about how best to maximize one's utility, so that if they didn't affect them in that way one wouldn't persist in the unhealthy behaviour? Note that Elster's objection to Becker – 'how can he tell the difference?' – cuts both ways. Addiction may act on belief formation as well as desire formation, but that doesn't tell us whether it would be rational for the addict to discontinue use if it didn't act on his beliefs in the ways it does.

Can T/S prove that their assumption #4 is true, and thus, that the objection of Glod, Shapiro, Gaus and Becker to it is mistaken? To the claim that the choices of smokers, overeaters, etc. are irrational, given

their beliefs and desires, it is always possible for the rational choice theorist to reply that the wrong beliefs and desires have been ascribed to them. Of course if the theorist needs to concoct some fairly exotic or *ad hoc* beliefs and desires to save her claim that their choices are utility maximizing, then the value of the theory as an explanatory theory in the social sciences becomes suspect. The standard reply to this worry by the theory's defenders is that its simplicity, elegance and explanatory power – together with the absence of any competing theory of comparable simplicity, elegance and power – justifies the occasional *ad hoc* hypothesis to handle otherwise disconfirming evidence.⁴

2. MULTIPLE SENSES OF 'ONE'S GOOD' AND TYPES OF PATERNALISM; BENEFICENCE AND DOCG; RATIONALITY AND PRUDENCE

The question whether most smokers, junk food bingers, helmetless motorcycle riders, spendthrifts, crack users, and so on are making irrational choices seems to me difficult to settle definitively one way or the other. I now wish to argue that even if Glod, Shapiro, Gaus and Becker are right and most of the people who engage in those behaviours do so rationally, that wouldn't undercut the case for an 'ends/prudential' paternalism that targets them. T/S's assumption #4 is not needed to make the case for that kind of paternalism.

Paternalism may be defined as 'interference' (via coercion/restrictions of liberty, deception, nudges or withholding information) with another person's choices or decision making for 'his own good'. The term 'one's good' is ambiguous and can refer to: (i) one's prudential good as identified by the correct normative theory of prudential value; (ii) one's (current) conception of one's prudential good (which may be mistaken according to the correct normative theory); and (iii) one's (current) 'conception of the good', one part of which is (ii). This ambiguity creates the possibility of different types of paternalism.⁵ *Prudential* paternalism aims at promoting (i), 'one's good' (welfare, well-being, self-interest) as identified by the correct theory of prudential value. (Paternalism with a different aim is non-prudential.) There are a number of competing theories of the prudential good for normal human beings, including hedonism, preferentism, objective list theories, and hybrid theories that combine elements of any two or all three of these (Parfit 1984; Griffin 1986).

⁴ See Hausman (1992), esp. chapter 13 for criticism of this reply by economists to the evidence of preference reversals provided by social psychologists. Røgeberg (2004) presses the objection of *ad hoc*-ness and unfalsifiability to rational addiction theory in particular.

⁵ The requirement that paternalism of any sort be aimed at promoting the 'good' of its target in some sense distinguishes paternalism from moralism. An example of the latter is a ban on 'same sex sodomy' for the reason that the conservative Christian favours one. It does not seek to promote the 'good' of gays/lesbians in *any* sense.

According to an objective list theory, some things (e.g. friendship, knowledge, artistic creativity and aesthetic appreciation) are intrinsically good for you regardless of whether you enjoy or have any pro-attitude toward them. The hedonist theory identifies one's self-interest with pleasure/the absence of pain, and the preferentist theory identifies it with the satisfaction of (some proper subset of) one's preferences weighted by their strength or ranking. *All* of these theories identify one's prudential good with what is most prudentially advantageous *over the long run*. Thus, according to the hedonist theory, one's self-interest is the maximization of one's present *and future* pleasure minus one's present *and future* pain (where both are measured cardinally by their intensity and duration). Beneficence is the moral duty (*prima facie*) that we have to promote the prudential good of others. It is often what motivates and/or justifies prudential paternalism.

One's 'conception of the good' (hereafter, CoG) is, roughly, the sum of one's present preferences, values or ends, ranked by the importance one attaches to them. (This is rough, because it ignores the important role that metaphysical beliefs and commitments might play in one's CoG.) Formulating and acting on a CoG requires a capacity for rational agency, while having a prudential good only requires a capacity for sentience. Toddlers and severely cognitively disabled adults have a prudential good but not a CoG. If one is a pure egoist, one's CoG just *is* one's conception of one's prudential good. But few are pure egoists; most people value intrinsically other things besides their own welfare. Many have altruistic preferences regarding some other people or animals, valuing *their* well-being for its own sake. And many hold values unrelated to anyone's welfare. Many conservationists believe that the preservation of species threatened with extinction by human activity has value whether or not it benefits our own species in any way, while conservative Christians believe that homosexuality is 'unnatural' and forbidden by God, making it wrong even when it is private conduct between consenting adults that neither harms nor wrongs anyone. It is true that whenever we act we are trying to satisfy our preferences, but it is not true that whenever we act we seek our own welfare/prudential good either as identified by the correct normative theory or as we conceive it. This is the reason why preferentism as a normative theory of prudential value, to be remotely plausible, cannot identify one's welfare simply with the satisfaction of one's preferences, but instead must screen out some types of preferences.⁶

⁶ Overvold (1980) argues that if it doesn't, it will make self-sacrifice impossible. (Prudential self-sacrifice *is* possible; psychological egoism is false.) Parfit (1984: 494) provides a counterexample to the 'unrestricted desire fulfillment' theory of welfare: suppose that you meet a man on a train who is believed to have a fatal illness; you strongly want him to be cured; you never meet him again or learn that he has been cured; it is absurd to think

Another distinction that we'll need is between what Sunstein calls 'means' vs. 'ends' paternalism.⁷ Means paternalism enables its targets to better realize their own CoGs or totality of ends (i.e. their actual, current ends – not the ends they *would* have under certain hypothetical circumstances, nor the ends they lack now but will probably have in the future). The means paternalist takes no stand on whether her target's ends are foolish or misguided, and instead justifies her interference solely on the grounds that it will enable him better to achieve his ends than his non-interfered-with choices would. Ends paternalism, by contrast, overrides its target's ends or values; the ends paternalist supposes that they are erroneous or defective in some way. So understood, the distinction between means and ends paternalism is exclusive.

The two paternalism distinctions we've drawn – means/ends and prudential/non-prudential – are independent of each other. Paternalism is prudential or not depending on the type of 'good' it seeks to promote, while it is means or ends depending on the ends/preferences of its target. The very same interference with choice – say, a nudge intended to discourage smoking – can be means paternalism when it targets one smoker but ends paternalism when it targets another, because of the different preferences/ends that the two smokers have. The independence of the two distinctions gives rise to four types of paternalism: means/prudential; means/non-prudential; ends/prudential; and ends/non-prudential. The first three types are the only ones that will figure in subsequent discussion, so the fourth type, ends/non-prudential, may be ignored. Note, however, that it is a possible type of paternalism only if there are other senses of 'one's good' besides the three distinguished above. Maybe we have an alethic or epistemic good in knowing certain truths whether or not knowing them benefits us. If so, then 'epistemic paternalism' (see Goldman 1991) is ends/non-prudential paternalism when it targets those with a CoG that embodies defective epistemic values (e.g. evangelical Christians who deny evolution).

The aim of the rational person is to advance her CoG. If one's CoG is (roughly) the sum of one's current ends ranked by the importance one attaches to them, then another way to say this is that the rational person seeks to maximize the satisfaction of her weighted, presently held preferences, *given her beliefs about the world*. This is 'economic rationality'.

that his recovery increases your welfare. Sumner (1996: 134–5) infers from Parfit's example that a preferential account of welfare must exclude 'disinterested' desires, while Hausman (2012: 81–4) infers from it that 'altruistic' (as well as 'malevolent') desires must be excluded.

⁷ Dworkin (2017) has the same distinction but with the different nomenclature of 'weak'/'strong'. Their distinction should not be confused with either of two 'soft'/'hard' distinctions described in footnote 14 below.

A slightly thicker conception of rationality adds the second requirement that those beliefs be epistemically reasonable, *given one's evidence or information*. A problem with the view that an Econ is nearly omniscient, having acquired all information that might be relevant to determining which of her options is optimal, is that it ignores the fact that information acquisition has costs. We should follow Jon Elster, who identifies as a third condition for a choice to be fully rational that it be based on 'an optimal amount of information', that is, the amount produced by 'the optimal amount of time, energy, and money in gathering [it]' (Elster 1999: 35). A choice is fully rational, on Elster's view, just when it satisfies all three of these conditions.

This account of rationality is 'instrumental' or Humean: it imposes no constraints on the content of one's CoG or any of one's ends. Thus, suppose that someone genuinely believed that it would be best to sacrifice one of his limbs if doing so were necessary to save the life of a housefly. Keeping his limbs is part of his prudential good (something we may suppose he himself recognizes), but saving the fly's life has a *non-prudential* value that outweighs it, according to his CoG. The Humean account implies that it would be rational for him to make the prudential sacrifice, if necessary. 'Prudence', by contrast, imposes constraints on one's CoG, implying that it is unreasonable if it gives insufficient weight to one's prudential good. The choice to sacrifice a limb to save a housefly is certainly imprudent on the assumption that the value of a housefly's life is simply not that great. If someone really held this bizarre CoG but hesitated to make the sacrifice when the occasion required it because he had a weak will, then a nudge to help him overcome his irrationality might be justifiable on means paternalist grounds, but it would not be justifiable on prudential paternalist grounds.⁸ Thus, nudging him would be an example of means/non-prudential paternalism.

Liberal theories of justice require that we 'tolerate' many of the CoGs held by others, even if we reject them or judge them mistaken. Of course we are free to criticize them, but we may not 'interfere' with others' choices to act on their CoGs, or try to 'impose' on them our own CoGs. I'll dub this principle 'DOCG' (for Deference to Others' CoGs). It implies that it would be wrong to thwart the choice of a non-akratic fly-lover to sacrifice one of his limbs to save a fly. DOCG is a deontological side-constraint, either *prima facie* or absolute, that forbids ends/prudential paternalism.

⁸ I assume that paternalism of any kind requires an 'interference' of some sort. On that assumption, refraining from nudging our weak-willed fly-lover is not ends paternalism even if it's an intentional 'omission' motivated by the belief that his ends are foolish. If omissions can be paternalistic – as argued in Shiffrin (2000: 213) – then the prudential paternalist's intentional non-interference makes her an ends paternalist in this case.

A few more comments in clarification of DOCG are in order. First, we certainly have no duty to allow white supremacists to act on their CoGs and oppress minorities, or defer to the wishes of conservative Christians to deny to gays/lesbians the benefits of legal marriage. This means that DOCG's scope has to be qualified in something like the following way:

For 'interference' with another's choices or choice making to be justified, it must be justifiable *either* by appeal to principles of justice and the public good to which all morally reasonable persons would agree, or by appeal to its targets' very own CoGs.

We may thwart the efforts of racists to oppress minorities, but not force them to socialize with minorities in their own homes. We may ignore the wishes of conservative Christians to deny the benefits of legal marriage to gays/lesbians, but we may not forbid them to worship in churches that teach that unrepentant homosexuals are eternally damned. Of course this 'qualification' of DOCG presupposes a great deal of liberal moral and political theory, here unspecified, about what the 'principles of justice and the public good to which all morally reasonable would agree' are, but that can't be helped. Arguably, the best specification of that phrase yields a DOCG principle that forbids not just ends/prudential paternalism but moralism (of the sort exemplified by the ban on 'same sex sodomy' favoured by the conservative Christian) as well. Second, whereas beneficence is a 'positive' duty, DOCG is only a 'negative' one. The duty is not to interfere with others' choices in pursuit of their currently held CoGs; it isn't to help them in their pursuits. The devout Catholic violates DOCG by trying to silence the defenders of 'false faiths', but not by refusing to aid them in their proselytizing efforts. Third, as a consequence of this last point, DOCG is supposed to give everyone a reason to refrain from ends/prudential paternalism, but it does not give anyone a reason to engage in means paternalism. It is a reason to refrain from interfering with the choice of the non-akratic fly-lover to make the sacrifice required by his CoG, not a reason to apply the nudge to the akratic fly-lover. Someone who shares the fly-lover's CoG has a subjective reason to nudge him, but neither beneficence nor DOCG gives anyone an objective reason to do so.

I shall say that a person is 'prudent' if: (a) her CoG does not misidentify or undervalue her own prudential good; (b) she has a temporally neutral concern for her prudential good (harms and benefits in the future don't count for less than equivalent ones in the present merely because of their temporal location, independent of how much that affects the likelihood of their occurrence); and (c) she is not excessively risk seeking or risk avoiding.⁹ Notice that the Humean account of rationality

⁹ Brink (2003), following Adam Smith and Henry Sidgwick, sees the first two features, proper weighing of one's prudential good and temporal neutrality, as essential to

does not require any of these. Satisfying the three conditions for rationality described above does not guarantee that one will have any concern for one's prudential good, much less a temporally neutral one that is not excessively risk seeking or risk averse.¹⁰ The rationality of risk-averse or risk-seeking choices all depends on one's attitude toward risk. If one is sufficiently risk seeking, it is rational to play lotteries with only very long odds of winning.

The 'prudent' person, defined as someone whose stable preferences include (a)–(c), will make choices that are not in his best interests if he holds false but reasonable beliefs about the long-term consequences of his choices. Even if such choices are 'prudent', if he can't be dissuaded from making them via rational persuasion (perhaps he not unreasonably doubts the credibility of those warning him), then prudential paternalism toward him becomes an option. Thus, prudential paternalism is possible toward the 'prudent'. Moreover, since his conception of his prudential good agrees with whatever the correct normative theory of prudential value requires, this prudential paternalism will also be a means paternalism. Means/prudential paternalism is possible, but only when its targets have preferences consistent with (a)–(c) above.

While the claim that a particular choice is both prudent and irrational comes close to being an oxymoron, the claim that a prudent person occasionally makes irrational choices does not. If 'prudent person' is defined solely in terms of the possession of (a)–(c), then he can make irrational choices due to weakness of will, fallacies of causal reasoning, etc. And if prudent persons can make irrational choices, then there will be even more opportunities for means/prudential paternalism toward them.

Whether or not Glod's overeaters, Shapiro's smokers, Gaus's helmetless motorcyclists and Becker/Murphy's hard drug users are irrational and/or imprudent, it seems pretty clear that at least for *most* of them, their behaviours are contrary to their long-term best interests.¹¹ This is so even if the correct theory of prudential value is a preferentist one. A strong desire shared by nearly all of us is for 'happiness' conceived

prudence, but he does not mention this third one concerning risk. The claim that it is another feature of prudence seems to me at least consistent with if not required by our commonsense understanding of that 'virtue'.

¹⁰ Hume himself is explicit about this: "Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. 'Tis not contrary to reason for me to chuse my total ruin, to prevent the least uneasiness of an Indian or person wholly unknown to me. 'Tis as little contrary to reason to prefer even my own acknowledge'd lesser good to my greater, and have a more ardent affection for the former than the latter' Hume (2007: 267).

¹¹ The 'most' qualification is necessary. The prudential paternalist who targets a group – smokers, overeaters, under-savers, etc. – has to admit that the behaviour she aims to discourage might in fact promote the long-term prudential good of a few, atypical members of the group.

in fairly hedonistic terms. Smokers and overeaters risk long periods of declining health and happiness before they die, during which time many of them will rue their unhealthy lifestyle choices. The expected value of that frustration and unhappiness, together with the expected value of the satisfaction and happiness they forego due to reduced health/longevity, probably outweighs any satisfaction and happiness that their smoking or overeating provides them. For the latter to outweigh the former, it would have to be the case that there are no other, more healthful 'stress relievers' available that are nearly as effective, or the costs of transitioning to them are extremely high. While that might be true for some smokers and overeaters, it is hard to believe that it's true for the majority of them.

Perhaps the defence of a prudential paternalism that targets the behaviours T/S would nudge is not as easy as has just been suggested. Nevertheless, it is not threatened by Glod's observation that many with unhealthy diets 'really enjoy eating and are willing to pay the price', by Shapiro and Gaus's claims that most smokers and helmetless motorcyclists are less risk averse than others, or Becker/Murphy's claim that hard drug users are higher future utility discounters than others. For even if they are right, that at most shows that the choices in question are not irrational. It does not show that they promote overall present and future welfare. If the choices/behaviours in question are rational, then interference with them cannot be defended on means/prudential paternalist grounds, but it can still be defended on ends/prudential paternalist grounds.

3. WHY T/S'S NUDGES ARE NOT MEANS/PRUDENTIAL PATERNALISM

However, T/S do not wish to defend their nudges on the latter grounds. Sunstein explicitly claims that the paternalism that he and Thaler support is the means type (Sunstein 2014: 75). Being means paternalism would explain why T/S suppose it is important to show that the behaviours their nudges are meant to discourage are irrational. For means paternalism permits interference with our choices only when they are inapt given our ends, and they will often be inapt if they are irrational.¹² T/S also say that the point of their nudges is to increase people's 'welfare' by leaving them 'better off as judged by themselves'. That sounds like prudential paternalism. Thus, the paternalism that they're defending is *supposed* to be both means and prudential.

However, it is neither. The reason why it isn't prudential paternalism is that its aim, to leave people 'better off as judged by themselves', is incompatible with any normative theory of prudential value. One's

¹² Often, not necessarily, because some irrational choices can be utility maximizing through dumb luck. There is no means paternalist reason to interfere with those choices.

judgement whether one is made better or worse off by some state of affairs reflects one's *current* conception of one's prudential good, which, unless one is 'prudent', will be mistaken because biased toward the present. If Sunstein had said that paternalism increases its targets' welfare only if it maximizes the satisfaction of their *present and future* conceptions of their prudential good, then he might have a coherent account of welfare, probably some form of a preferentist theory of prudential value.¹³ If that were the aim of his paternalism, then it would be prudential. But if that were its aim, then he would have to admit that it will sometimes override its target's current conception of his prudential good for the sake of conceptions the target is likely to hold in the future, which would disqualify it from being means paternalism.

The rejection of ends paternalism by T/S's libertarian critics (including Glod, Gaus and Shapiro) is due to their acceptance of DOCG as a deontological side-constraint.¹⁴ If the Sunstein of *Why Nudge?* is a welfarist consequentialist of some sort, as he appears to be (Sunstein 2014: 132–6), then a deontological DOCG cannot be his reason for rejecting it. Although he doesn't explain why he rejects it, he might do so for the following reason. Ends/prudential paternalism will sometimes override its target's current conception of his prudential good, leaving him worse off as judged by himself. Since Sunstein assumes that paternalism increases its targets' welfare only if it leaves them better off as judged by themselves, he might reject ends paternalism on the grounds that it is necessarily ineffectual; it cannot possibly promote its targets' welfare. The problem with this reasoning lies with Sunstein's assumption, which confuses one's prudential good with one's current conception of one's prudential good.

For T/S's nudges to be a means/prudential paternalism based on a coherent preferentist account of welfare (one that screens out certain types of preference and identifies one's prudential good with maximum satisfaction of the remaining present *and future* preferences), it would

¹³ Note, however, that 'one's prudential good consists in maximizing the satisfaction of one's present and future conceptions of one's prudential good' is viciously circular. A preferentist account of the good that tries to solve the problem of specifying which preferences are excluded (the issue broached in footnote 6 above) by saying that it is only preferences for one's own prudential good that count suffers from the same vicious circularity, as Sumner (1996: 135) notes.

¹⁴ Hodson (1977) and Vandever (1986) also seem to reject ends/prudential paternalism on the basis of DOCG (or something akin to it). Joel Feinberg rejects what he calls 'hard' paternalism (which for him is paternalism that interferes with 'substantially voluntary' choices, not – as for T/S and many economists – paternalism that uses coercive means) because it 'imposes its own values and judgments on people' (Feinberg 1986: 12), suggesting that he too accepts DOCG. An objection to Feinberg is that while ends paternalism 'imposes its own ends' on its targets, 'hard' paternalism as he defines it needn't. For more on this objection, see Scoccia (2008: 354–7).

have to be the case that its targets are 'prudent' persons as defined in the previous section. If they have the preferences of the prudent but they make irrational choices, then T/S's nudges would elicit choices that better satisfy their true, context independent preferences for good health, a financially secure retirement, a long life, and so on.

T/S's libertarian critics think it obvious that many of the smokers, overeaters, etc. targeted by T/S's nudges do *not* have the preferences of the prudent person, and thus, that those nudges are in fact an ends paternalism forbidden by DOCG. Till Grune-Yanoff objects that T/S's paternalism respects 'neither the subjectivity nor the plurality of people's values' (Grune-Yanoff 2012: 641), while Gregory Mitchell objects that it seeks to 'impose some conception of objectively good preferences on irrational persons' (Mitchell 2004: 1260).¹⁵ These critics seem to me correct. Many smokers, overeaters and so on avow and act on a conception of their prudential good in the neighbourhood of 'eat, drink, and be merry, for tomorrow we die' that's incompatible with the prudent person's preferences. The fact that many of them regret their imprudent choices 'when the chickens come home to roost' does not show that they really had a temporally neutral and not excessively risk-seeking concern for their welfare when they earlier engaged in the imprudent behaviours.

Feinberg (1986: 109) notes that there are some who regard 'niggling prudence' as 'dull and unappealing ... better the life of spontaneity, impulse, excitement, and risk, even if it be short, and even if the future self may bear the costs'. The prudent in their view are timid, boring Caspar Milquetoast-types. Some of them may not regret their earlier lifestyle choices at all but instead view the later prudential harms those choices led to as the price they must pay for the good life they've lived.¹⁶ Three points should be made about them. First, it is *possible* that at least for

¹⁵ If Mitchell means that the paternalism T/S support is *in fact* ends paternalism, then he's correct. But if he means that T/S *intend* to support an ends/prudential paternalism (i.e. impose some conception of 'objectively good preferences' on persons whom they admit don't hold those preferences), then he is incorrect.

Mitchell has also objected that 'libertarian paternalism' is an 'oxymoron'. I think that too is incorrect. The libertarian position on the permissibility of any kind of state paternalism seems to me to be: (i) ends/prudential paternalism is forbidden whether it uses coercive shoves or non-coercive nudges; and (ii) means paternalism is permitted, but only if it uses non-coercive nudges. Note that if the libertarian accepts (ii), he will have to point to some rationale other than DOCG to explain why coercion in the service of means paternalism is forbidden, since DOCG does not forbid coercion in that circumstance.

¹⁶ Christopher Hitchens' obituary in the *New York Times* (16 December 2011) reports him as having said: 'Writing is what's important to me, and anything that helps me do that – or enhances and prolongs and deepens and sometimes intensifies argument and conversation – is worth it to me.' He found it 'impossible to imagine having my life without going to those parties, without having those late nights, without that second bottle'.

some of them their earlier lifestyle choices were *not* contrary to their long-term prudential good on the preferentist account of it. The prudential paternalist can admit as much but must claim that they are atypical of smokers, drug/alcohol abusers, under-savers, etc. in general. Second, a sincere, defiant insistence that one has 'no regrets' is not proof positive that one's earlier choices did promote one's prudential good. It may just reflect a combination of stubbornness, self-deception and the 'sunk costs' fallacy. Finally, and most importantly, even if the earlier lifestyle choices that led to the later prudential harms are imprudent, a prudential paternalism that targets them remains an ends paternalism that violates DOCG.

While T/S's libertarian critics might dispute my contention at the end of the last section that the lifestyle behaviours targeted by T/S's nudges are prudentially harmful for most of those who engage in them, they needn't. Their claim is that ends/prudential paternalism is wrong *even if* the choices made by its targets *really are* imprudent, because it 'imposes' on those targets values that they reject, *contra* DOCG. The value it seeks to impose is that of *prudence itself*.¹⁷ It would impose on Shapiro's smokers, Glod's overeaters and Gaus's helmetless motorcycle riders a 'you should care about your happiness in a more temporally neutral and less risk seeking way' value that is foreign to their own current conception of their prudential good (leaving them 'worse off as judged by themselves'). This 'imposing' is not as forceful if it uses nudges rather than coercive shoves, but it is 'imposing' in either case.

An ends/prudential paternalism based on a narrow perfectionist or objective list account of the good will probably support even more 'imposing of values' than one based on a hedonist or preferentist account. The former might support forcing people who spend their leisure time getting high and watching reality TV shows to study philosophy or read *Anna Karenina* instead. But because prudence demands an attitude toward one's future welfare that's inconsistent with the strong bias toward the present that many people have, even a prudential paternalism based on a preferentist account of welfare will end up being an ends paternalism for many of its targets. The defence of an ends/prudential paternalism that targets smoking, overeating, under-saving, and so on does *not* presuppose the (absurd) objective list theory that includes health and longevity on its

¹⁷ Note that of the three conditions for full rationality mentioned earlier, the first, 'economic rationality', involves a purely 'formal' end that everyone necessarily has, and thus, unlike prudence, is not a value that can be 'imposed' on anyone. The same probably cannot be said for the second condition, epistemic reasonableness. Thus, Shapiro (1994: 191) notes: 'When the irrationality that smokers manifest is due to epistemic values that are an important part of their conception of the good, then it will be untrue that making it more difficult for irrational smokers to smoke simply makes it easier for them to achieve their ends.'

list and claims that the slightest gains in them outweigh the largest losses in other goods (what Crampton (2009) calls 'healthism').

4. AUTONOMY OBJECTIONS TO T/S'S NUDGES; SUNSTEIN'S REPLY

There is another reason why the nudges that T/S support will turn out to be ends paternalism with regard to many of their targets. Many people seem to value 'autonomy' as a non-prudential good greatly and for that reason would oppose paternalistic interference that limits it, even if the interference would significantly benefit them in the long run and they value their welfare in a temporally neutral way. This fact gives rise to an 'autonomy objection' to T/S's nudges. In fact, I will argue that there are two very different objections alleging that their paternalism would 'violate autonomy', one Kantian and the other non-Kantian. The latter is based on the DOCG principle. While Sunstein does have a reply to it, his reply is unconvincing.

The DOCG principle is one way of understanding the duty to 'respect others' autonomy'. DOCG demands deference to others' CoGs, and respect for the autonomy of others plausibly requires deference to their CoGs as well as non-interference with their choices. Consider a case in which a man has suffered a serious injury, a blood transfusion is medically indicated, but he is unconscious. There is a clear sense in which we 'violate his autonomy' if we give him the transfusion despite knowing that he's a Jehovah's Witness who would not consent to it if he were conscious. The reason it 'violates his autonomy' is not that it would interfere with any of his choices (he can't make any in his current condition, and he may never have previously, explicitly indicated how he would want others to treat him in the present circumstances) but rather that it fails to defer to his CoG.¹⁸

Now a feature of the CoGs held by many people is disapproval of unauthorized interference with one's choice making by others, even if the interference resulted in better choices or outcomes as judged from one's own point of view. Of course we are happy to delegate some of our choice making to others, especially experts with a fiduciary responsibility to look out for our best interests (e.g. an accountant to complete one's taxes). We are also okay with unauthorized interference in certain extreme cases. If a normal person walking down the street is distracted by texts on his mobile phone and about to step in the path of a quiet, speeding electric tram, he would surely welcome a good Samaritan's lifesaving shove, even if he had never previously consented to it. The normal person

¹⁸ We might also say that it violates his autonomy because he *would* not consent to the transfusion if he *were* conscious and able to make a choice. There are different ways to specify a 'hypothetical consent' test for the permissibility of paternalistic interference, at least one of which is equivalent to DOCG.

values the avoidance of unforeseen, imminent and violent death more than being allowed to act on his own choices. But in cases where a choice is likely to produce only slight prudential harm (e.g. a mild hangover from consuming too much alcohol), many of us would prefer to suffer the harm to suffering an interference not previously consented to that would prevent it ('I agree that I will be better off if I don't have a third martini, but I prefer being allowed to make that slightly foolish choice to others' thwarting it for my own good'). Autonomy is being valued in such cases as a non-prudential good, tied to one's sense of oneself as a person entitled to others' respect. Someone who would interfere with my slightly foolish choices without my consent for my own good does not treat me with respect, but instead treats me 'like a child'. Most of us deeply resent being treated as if we were incompetent. The non-Kantian autonomy objection to T/S's nudges alleges that they will often be ends paternalism in violation of DOCG, because many of their targets strongly value autonomy as a non-prudential good. The nudges will often 'violate autonomy' for this reason, as well the reason described in the previous section (i.e. by 'imposing' substantive values like health, longevity, physical safety, long-term financial security, etc. on people who don't value them as strongly as T/S claim they do).

Sunstein has a reply to this objection, but before considering it we should note how the autonomy objection just described differs from the Kantian autonomy objection. The main difference is that the non-Kantian one judges unauthorized interference with one's choice-making (via coercion, deception, withholding information, nudging) acceptable if it doesn't 'impose' on one any values alien to one's CoG, whereas the Kantian objection judges it wrong even if one does not value autonomy (for its own sake and as a non-prudential good) at all. In other words, the non-Kantian objection condemns unauthorized coercion, deception, etc. only when they are used in ends paternalism, not when used in means paternalism. (I'm assuming that if someone takes inapt means to her end X, but she also values autonomy more than X, then nudging her to make the choice that will realize X is ends paternalism, *not* means paternalism, because it ignores the fact that her CoG prizes autonomy more than X.) The Kantian objection condemns unauthorized coercion, deception, etc. *per se*, whether it is used in means *or* ends paternalism.

The Kantian objection is based on the Principle of Humanity, the formula of the Categorical Imperative requiring that one treat 'humanity' both in oneself and others always as an 'end' and never as a 'mere means'. It assumes that the exercise of one's capacity for practical rationality is an intrinsic non-prudential value that *trumps* both one's welfare and the pursuit of any other ends one might have. It doesn't matter whether the paternalist's target is indifferent to autonomy as an end, valuing it merely as a means to achieve his other ends. The person who would

not indignantly condemn a nudge or shove that he hadn't previously authorized, but that enabled him better to achieve his ends, is deficient in self-respect; he treats himself and his personhood/rational agency capacities as 'mere means'.¹⁹ The assumption here is that to respect others as persons requires that one try to influence their decision making only through transparent rational persuasion, not any of the underhanded means that paternalists use to control their targets' choices. This Kantian version of the 'autonomy objection' to paternalism underlies the 'nudging is manipulative and manipulation is intrinsically wrong' objection to T/S's assumption #5 noted earlier.

I assume that the Kantian objection isn't meant to apply to unauthorized interference in cases where 'impairment' precludes the exercise of practical rationality/rational agency. For example, if you choose *while drunk* to accept a dare to perform some very dangerous stunt, I do not 'violate your autonomy' if I thwart your choice. Presumably strong depression and extreme fatigue, anger or fear also count as 'impairments'. The Kantian view permits mandatory 'cooling off' periods, especially for important and irrevocable choices, that prevent one from making any choice till one has regained full use of one's rational faculties. But the Kantian view faces a problem when we ask what exactly is supposed to count as 'impairment'. If the distraction in the mobile phone/speeding tram example doesn't count as one, then the Kantian view absurdly forbids the good Samaritan's shove in that example. If it does count, then why not factually mistaken belief about which means will achieve one's ends, impulsiveness, a weak will, cognitive bias, and so on? But if the bar for what counts as 'impairment' is set that low, then the Kantian view will forbid no means paternalism at all. Yet forbidding unauthorized interference in means paternalism cases was supposed to be what distinguishes that view from the non-Kantian version of the autonomy objection.

My concern is not to pass final judgement on whether the Kantian objection has merit, but to emphasize that it is a different objection from the non-Kantian one. Hausman and Welch (2010) appear to conflate them when they say:

[I]n addition to or apart from rational persuasion, [nudges] may 'push' individuals to make one choice rather than another. Their freedom, in the sense of what alternatives can be chosen, is virtually unaffected, but when this 'pushing' does not take the form of rational persuasion, their autonomy – the extent to which they have control over their own evaluations and deliberation – is diminished. Their actions reflect the tactics of the

¹⁹ On the Kantian view of the connection between self-respect and valuing one's capacity for rational agency and choice autonomy over one's own welfare, see Hill (1973).

choice architect rather than exclusively their own evaluation of alternatives. Hausman and Welch (2010: 128)

The claim that nudges ‘diminish autonomy’ because they are not ‘rational persuasion’ is the Kantian objection, while the suggestion that T/S’s nudges do not accord with their targets’ ‘own evaluation of alternatives’ is the non-Kantian one.

Let’s return to the non-Kantian objection and consider Sunstein’s reply to it. Sunstein draws a useful distinction between ‘thick’ and ‘thin’ conceptions of autonomy. According to the ‘thick’ one, autonomy is ‘an end in itself ... To treat people with respect, and as ends rather than mere means, government cannot override that form of freedom even if doing so would, in fact, make people happier or better off in a relevant sense’ Sunstein (2014: 127). Thus, Sunstein’s thick conception implies that trade-offs between autonomy and welfare are possible. According to his thin conception there can be no such trade-offs, only trade-offs between the autonomy component of one’s welfare and other components of it. It follows that the thick conception makes autonomy a non-prudential value, while the thin one supposes that it is valuable only as a means to or a component of one’s prudential good.²⁰

Having drawn this distinction, Sunstein proceeds to argue that people do *not* in fact value autonomy in its thick, non-prudential form. Some may *think* that they value it in that way, but they *really* don’t. He says:

What really does and should matter is welfare, for which claims about autonomy are best understood as a heuristic ... Autonomy is what matters to System 1, but on reflection, the real concern, vindicated by System 2, is welfare. People speak in terms of autonomy, but what they are doing is making a rapid, intuitive judgment about welfare. (Sunstein 2014: 134)

What are we to make of this? Does Sunstein’s claim about how people value autonomy refute the non-Kantian autonomy objection?

It would if it were correct. If people did in fact value autonomy only as a prudential good – intrinsic, instrumental, or both – then it would make no sense for them to oppose a paternalism that leaves them all things considered better off. For even if it reduces their autonomy, its doing so is a prudential cost that must be outweighed by its prudential benefits, otherwise it wouldn’t be a paternalism that leaves them all things considered better off.

But Sunstein’s claim that few if any value autonomy as a non-prudential good is no more credible than his claim that smokers, overeaters and so on make irrational choices that are contrary to their

²⁰ See Feinberg’s discussion of ‘four standard ways of treating the relation between personal autonomy and personal good’ (Feinberg 1986: 58). Sunstein’s thin conception of autonomy is the first one, which Feinberg rejects.

own deepest preference to seek their prudential good in a temporally and risk neutral way. Consider those smokers who do *not* believe that the 'pleasures of smoking' now outweigh the reduced health/longevity that their habit will lead to years later. They admit (now) that smoking is bad for them and they ought to quit, but they have weak wills. They are prudent persons making an irrational choice. No doubt some of them would welcome an intervention that allows them to avoid the consequences of their own irrationality. In their case it would be means paternalism. But surely many others, because they value autonomy as something crucial to their self-respect more than they want to avoid the likely consequences of their irrational and imprudent choice to continue smoking, would not welcome it. A prudential paternalism that targets them is ends paternalism.

5. A BETTER DEFENCE OF PRUDENTIAL PATERNALISM?

Sunstein's reply to the non-Kantian autonomy objection is the only one he can give so long as he holds that his public health nudges are permissible only if they are means paternalism. But that reply is unsatisfactory, and it is a reason (along with the one described in [Section 3](#)) why, if he wishes to defend those nudges, he has to abandon that view. He has to admit that his paternalism is a prudential paternalism that will often be ends paternalism. And that entails rejecting DOCG as a deontological side-constraint.

I don't mean to suggest that a defence of prudential paternalism as an ends paternalism requires a rejection of all deontology in favour of welfarist consequentialism (one that avoids Sunstein's mistake about welfare). One can consistently reject DOCG not just as an absolute but also as a *prima facie* deontological requirement without rejecting *all* deontological side-constraints. What I do suggest is that a defence of T/S's nudges requires the admission that, given the true preferences had by many of the targets of those nudges, it will often be ends/prudential paternalism. Thus, it requires a rejection of at least DOCG.

The prudential paternalist who rejects DOCG will claim that there is nothing intrinsically wrong with thwarting choices based on CoGs that embody false judgements about the good. (That doesn't mean that it is always permissible to thwart them; often there will be good consequentialist reasons not to.) DOCG, she must argue, depends for its plausibility on an untenable subjectivism or scepticism about the good. Her reply to the objection that T/S's nudges wrongly impose the value of prudence on those whose conception of their prudential good reflects a strong bias toward the present is that their conception is *mistaken*, and its being mistaken cancels out whatever *prima facie* reason there was to defer to it. According to this reply, Humean rationality may not require

that one have the aims and values of the 'prudent' person, but a richer more accurate account of practical rationality (perhaps Aristotelian) does. Certainly an imprudent choice by someone committed to a high rate of pure time discounting is unreasonable if Derek Parfit (and before him, Henry Sidgwick) are right in thinking that practical reason requires a temporally neutral weighing of costs and benefits. The Humean view is compatible with temporal preferences that seem irrational on their face.²¹

This prudential paternalist can also admit what Sunstein refuses to about autonomy: many smokers value it as a non-prudential good and to a degree that makes it rational for them to oppose paternalistic interference with their irrational and imprudent choice not to quit. But she will claim that these smokers are simply *mistaken* to value it in that way. It is reasonable to fear that the state might use nudges to promote the agendas of special interest groups *at the expense of one's welfare*. But even if autonomy has some non-prudential value, it does not have so much as to make it reasonable to oppose state nudges that dramatically *would* increase one's long-term welfare.

Whether this defence of prudential paternalism is an improvement over Sunstein's is a difficult question that I make no attempt to answer here. Defenders of DOCG will certainly deny that it is. They may protest that they are not sceptics about the good at all; they can and do admit that an ends/prudential paternalism that targets myopic smokers, overeaters, etc. really would be prudentially best for them, notwithstanding their beliefs to the contrary. But the fact that it would be best for them does not make it right, because the duty to respect their autonomy overrides the duty of beneficence. Again, I take no stand here on whether this view is correct. I claim that Sunstein's own defence of his public health paternalism as an exclusively means paternalism is a non-starter, and thus, if the defence of prudential paternalism sketched in the last couple of paragraphs is no better, then the nudges that T/S call for are probably not justified at all.

ACKNOWLEDGEMENTS

The author thanks Tim Cleveland, Jean Paul Vessel, and Mark Walker for helpful comments on an earlier draft of this paper.

²¹ This is the point of Parfit's Future Tuesday Indifference example. Imagine a man cares about his future well-being, but he cares about pleasures and pains that occur on any future Tuesday less than ones that occur on any other future day. Clearly this man has an unreasonable preference. See Parfit (1984: 124–6).

REFERENCES

- Becker, G. and K. M. Murphy. 1988. A theory of rational addiction. *Journal of Political Economy* 96: 675–700.
- Bovens, L. 2008. The ethics of nudge. In *Preference Change: Approaches from Philosophy, Economics, and Psychology*, ed. T. Grüne-Yanoff and S. O. Hansson, 207–219. Dordrecht: Springer.
- Brink, D. 2003. Prudence and authenticity. *Philosophical Review* 112: 215–245.
- Crampton, E. 2009. Public health and the new paternalism. *Policy Magazine* 25: 36–40.
- Dworkin, G. 2017. *Paternalism*. In *Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta. <<https://plato.stanford.edu/archives/spr2017/entries/paternalism/>>
- Elster, J. 1999. Rationality and addiction. In *Drugs and the Limits of Liberalism: Moral and Legal Issues*, ed. Pablo De Graff, 25–45. Ithaca, NY: Cornell University Press.
- Feinberg, J. 1986. *Harm to Self*. Oxford: Oxford University Press.
- Gaus, G. 2003. Liberal neutrality: a compelling and radical principle. In *Perfectionism and Neutrality: Essays in Liberal Theory*, ed. S. Wall and G. Kosko, 137–166. Lanham, MD: Rowman and Littlefield.
- Glod, W. 2015. How nudges often fail. *Social Theory and Practice* 41: 599–617.
- Goldman, A. 1991. Epistemic paternalism: communication control in law and society. *Journal of Philosophy* 88: 113–131.
- Griffin, J. 1986. *Well Being*. Oxford: Oxford University Press.
- Grune-Yanoff, T. 2012. Old wine in new casks: libertarian paternalism still violates liberal principles. *Social Choice and Welfare* 38: 635–645.
- Hausman, D. 1992. *The Inexact and Separate Science of Economics*. Cambridge: Cambridge University Press.
- Hausman, D. 2012. *Preference, Value, Choice, and Welfare*. Cambridge: Cambridge University Press.
- Hausman, D. and B. Welch. 2010. Debate: to nudge or not to nudge. *Journal of Political Philosophy* 18: 123–136.
- Hill Jr., T. 1973. Servility and self-respect. *The Monist* 57: 87–104.
- Hodson, J. 1977. The principle of paternalism. *American Philosophical Quarterly* 14: 61–69.
- Hume, D. 2007. *A Treatise of Human Nature*. Oxford: Oxford University Press.
- Mitchell, G. 2004. Libertarian paternalism is an oxymoron. *Northwestern University Law Review* 99: 1245–1278.
- Overvold, M. 1980. Self-interest and the concept of self-sacrifice. *Canadian Journal of Philosophy* 10: 105–118.
- Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Rizzo, M. and D.G. Whitman 2009. Little Brother is watching you: new paternalism on the slippery slopes. *Arizona Law Review* 51: 684–739.
- Rogeberg, O. 2004. Taking absurd theories seriously: economics and the case of rational addiction theories. *Philosophy of Science* 71: 263–285.
- Scoccia, D. 2008. In defense of hard paternalism. *Law and Philosophy* 27: 351–381.
- Shapiro, D. 1994. Smoking tobacco: irrationality, addiction, and paternalism. *Public Affairs Quarterly* 8: 187–203.
- Shiffrin, S. 2000. Paternalism, unconscionability doctrine, and accommodation. *Philosophy and Public Affairs* 29: 205–251.
- Sugden, R. 2008. Why incoherent preferences do not justify paternalism. *Constitutional Political Economy* 19: 226–248.
- Sumner, L. W. 1996. *Welfare, Happiness, and Ethics*. Oxford: Oxford University Press.
- Sunstein, C. 2014. *Why Nudge? The Politics of Libertarian Paternalism*. New Haven, CT: Yale University Press.
- Thaler, R. and C. Sunstein. 2003. Libertarian paternalism is not an oxymoron. *University of Chicago Law Review* 70: 1159–1201.

- Thaler, R. and C. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New York, NY: Penguin Books.
- Vandeveer, D. 1986. *Paternalistic Intervention: The Moral Bounds of Beneficence*. Princeton, NJ: Princeton University Press.

BIOGRAPHICAL INFORMATION

Danny Scoccia is Professor Emeritus in the Philosophy department at New Mexico State University. He writes on topics in social/political philosophy and applied ethics.