



DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Do People Like Nudges?

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

Citation	Cass R. Sunstein, Do People Like Nudges? (May 8, 2015).
Accessed	May 25, 2018 12:45:57 AM EDT
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:16147874
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

(Article begins on next page)

Do People Like Nudges?

Cass R. Sunstein*

Abstract

In recent years, there has been a great deal of debate about the ethical questions associated with “nudges,” understood as approaches that steer people in certain directions while fully maintaining freedom of choice. Evidence about people’s views cannot resolve the ethical questions, but in democratic societies, those views will inevitably affect what governments are willing to do. Existing evidence remains preliminary, but it supports five general conclusions. First, there is widespread support for nudges, at least of the kind that democratic societies have adopted or seriously considered in the recent past. Second, that support diminishes when people fear that because of inertia and inattention, citizens might end up with outcomes that they do not like. Third, there appears to be somewhat greater support for nudges that appeal to conscious, deliberative thinking than for nudges that affect subconscious or unconscious processing, even though many people will approve of the latter as well (especially if they are meant to combat self-control problems). Fourth, people’s assessment of nudges in general will be greatly affected by the political valence of the particular nudges that they have in mind (or that are brought to their minds). Fifth, transparency about nudging will not, in general, reduce the effectiveness of nudges, because most nudges are already transparent, and because people will not, in general, rebel against nudges.

I. Introduction

The last several years have seen an outpouring of work on “nudges,” understood as interventions that steer people in particular directions but that also allow them to go their own way.¹ A reminder is a nudge; so is a warning. A GPS nudges; a default rule nudges. Disclosure of relevant information (about the risks of smoking or the costs of borrowing) counts as a nudge. Save More Tomorrow plans, allowing employees to sign

* Robert Walmsley University Professor, Harvard University. This essay is a revised and expanded version of a chapter in a book-in-progress, *The Ethics of Nudging*, tentatively scheduled for publication in late 2016. Special thanks to Arevik Avedian for help with the survey presented here. (Note to readers: It is expected that later versions of this preliminary draft will include additional surveys.)

¹ See Richard Thaler and Cass R. Sunstein, *Nudge* (2008); Richard Thaler, *Misbehaving* (2015); David Halpern, *The Nudge Unit* (forthcoming 2015); OECD, *Regulatory Policy and Behavioral Economics* (2014); Rhys Jones et al., *Changing Behaviours: The Rise of the Psychological State* (2013).

up to give some portion of their future earnings to pension programs, are nudges.² A recommendation is a nudge. A criminal penalty, a civil fine, and a subsidy are not nudges, because they impose significant material incentives on people's choices.³

In many nations, public officials have been drawn to nudges, especially in recent years.⁴ If governments can achieve policy goals with tools that do not impose high costs, and that preserve freedom of choice, there is good reason to take those tools seriously. In domains that include savings policy,⁵ environmental protection,⁶ and health care,⁷ among others, behaviorally informed approaches have attracted considerable attention, and often led to concrete reforms.⁸ At the same time, some people have raised serious ethical concerns and objections.⁹ An evident question is whether nudges should be counted as unacceptably manipulative or as an interference with freedom, rightly understood.¹⁰ To make progress on the ethical questions, we could refer to defining commitments of various kinds – involving autonomy, dignity, welfare, and self-government – and ask whether some, many, or all nudges run afoul of those commitments.¹¹

This is a normative task, not an empirical one. But while the normative discussion continues, it is worthwhile to ask some empirical questions. What do people actually think about nudging and choice architecture? Do they have serious ethical objections? Or

² See Thaler, *supra* note.

³ On some of the complexities here, see Cass R. Sunstein, *Why Nudge?* (2014).

⁴ See David Halpern, *The Nudge Unit* (forthcoming 2015); OECD, *Regulatory Policy and Behavioral Economics* (2014); Rhys Jones et al., *Changing Behaviours: The Rise of the Psychological State* (2013); Cass R. Sunstein, *Simpler* (2013); House of Lords, *Science and Technology Select Committee, Behavior Change* (2011).

⁵ See Thaler, *supra* note.

⁶ See *New Perspectives for Environmental Policies through Behavioral Economics* (Frank Beckenbach and Walter Kahlenborn eds., forthcoming 2015); *Handbook of Research on Sustainable Consumption* (Lucia Reisch and John Thøgersen eds. 2015).

⁷ Douglas Hough, *Irrationality in Health Care: What Behavioral Economics Reveals About What We Do and Why* (2014).

⁸ See Sunstein, *Simpler*, *supra* note.

⁹ The best discussion is Riccardo Rebonato, *Taking Liberties* (2012). See also the various contributions to Special Issue, *Review of Philosophy and Psychology* (forthcoming 2015); Jeremy Waldron, *It's All For Your Own Good*, *New York Review of Books* (2014), available at <http://www.nybooks.com/articles/archives/2014/oct/09/cass-sunstein-its-all-your-own-good/> Consider in particular this question: "Deeper even than this is a prickly concern about dignity. What becomes of the self-respect we invest in our own willed actions, flawed and misguided though they often are, when so many of our choices are manipulated to promote what someone else sees (perhaps rightly) as our best interest?"

¹⁰ T. M. Wilkinson, *Nudging and Manipulation*, 61 *Political Studies* 341 (2013).

¹¹ For a preliminary effort, see Cass R. Sunstein, *Ethical Nudging*, *Yale J Reg* (forthcoming 2015).

do they believe that nudges are acceptable or desirable, even morally obligatory? Do they distinguish among nudges? What kinds of distinctions do they make?

The answers cannot, of course, dispose of the ethical questions. The issue is how to resolve those questions in principle, and empirical findings about people's answers are not decisive. Perhaps those answers are confused, insufficiently considered, or wrong. Perhaps people do not value autonomy highly enough, or perhaps they do not quite know what the concept means. Perhaps different nations, and different groups within the same nation, offer different answers, suggesting an absence of consensus. Behavioral scientists would emphasize a related point: People's answers to ethical questions, or questions about moral approval or disapproval, might well depend on how such questions are framed; slight differences in framing can yield dramatically different answers. Those differences are themselves a nudge.

Here is a small example of how ethical judgments can depend on framing.¹² If you ask people whether they think that young people should be valued more than old people, they will usually say, "certainly not." They will strenuously resist the idea that government should give a higher value to young lives than to old ones. But suppose that you ask people whether they want either (1) to save 70 people under age of 5 or (2) to save 80 people over the age of 80. It is reasonable to speculate that most people will choose (1), thus demonstrating that they are willing to value a young person more than an old one.¹³ It would be child's play to frame nudges so as to elicit one's preferred answer to ethical questions.¹⁴

Notwithstanding these points, people's answers to carefully designed questions do help to illuminate ethical problems, and for two different reasons. The first and most important is that in a democratic society, it is inevitable that public officials will attend to what people actually think, and to the extent that self-government is one of our goals, it is also desirable that they do so. If citizens have strong ethical objections, democratic governments will and should hesitate before proceeding (if only because of electoral self-interest). No public official will entirely disregard a strongly felt moral objection on the part of significant segments of the public. And if people do not have such objections, and if they welcome nudges as helpful and desirable, public officials will be, and should be, attentive to their views.

The second reason is epistemic: People's judgments provide relevant information about to think about the ethical issues even if that information is not conclusive. It is not necessary to make strong claims about the wisdom of crowds, especially on contested ethical issues, in order to believe that an ethical judgment, on the part of those who might be subject to nudges, deserves respectful attention. Public officials should be humble and

¹² See Shane Frederick, *Measuring Intergenerational Time Preference: Are Future Lives Valued Less?*, 26 *J. Risk and Uncertainty* 39 (2003).

¹³ See *id.*

¹⁴ For evidence, see *id.*

attentive to the views of others, and if strong majorities favor or oppose nudges, then their views ought to be taken into account.

As we shall see, current research offers two central findings. The first is that people tend to be enthusiastic about nudges -- at least if they take the form of those that have been seriously proposed, or acted on, by actual institutions in recent years.¹⁵ The second is that people agree that the nature of the particular nudge in question is highly relevant to the ethics of nudging. They tend to prefer nudges that target deliberative processes to those that target unconscious or subconscious processes – though they do not by any means rule the latter out of bounds, and will often approve of them as well. Their evaluation of nudges much turns on their political valence, which suggests that it is often people’s assessment of the ends of particular nudges, rather than of nudging as such, that settles their judgments. If people focus on particular nudges that they think to be ill-motivated, intrusive, threatening, or otherwise objectionable, they are likely to oppose nudges as such, and if they focus on particular nudges of which they approve, their overall evaluation tends to be positive. These findings fit well with the conclusion that in general, people are prepared to be highly supportive of nudges as policy tools.

II. Thumbs Up or Thumbs Down?

As a general test, I used Amazon Mechanical Turk to survey 296 people, asking whether they approved or disapproved of five well-known nudges: (1) mandatory calorie labels at chain restaurants; (2) mandatory graphic warnings on cigarette packages; (3) a mandatory “traffic lights” system for food, with green, red, and yellow labels with respect to health; (4) automatic enrollment in savings plans, subject to opt out; and (5) automatic enrollment in “green energy” providers, subject to opt out. About 44 percent of the respondents were Democrats; 41 percent were independents; and 15 percent were Republicans.

All five nudges attracted strong majority support. Nearly 96 percent favored calorie labels; 71 percent favored graphic warnings; 57 percent approved of traffic lights; 77 percent approved of automatic enrollment in savings plans; and 75 percent favored automatic enrollment in green energy. Notably, and perhaps surprisingly, the differences across party identifications were quite modest. Democrats and independents showed essentially no differences. Republicans showed majority support for all five nudges, though with modestly lower levels than Democrats and independents; the largest disparity by far involved automatic enrollment in green energy (supported by 79 percent of Democrats, 78 percent of independents, but just 55 percent of Republicans). Interestingly, the traffic lights system, which received the lowest overall level of support,

¹⁵ As discussed below, we could easily imagine nudges that would not attract much support. As we shall also see, ethical judgments about nudges in general might well be a product of examples that readily come to mind (a point suggesting the possibility of research into the psychology of ethical reactions to nudges and other policy tools, such as cap-and-trade).

was not regarded much less favorably by Republicans (59 percent support by Democrats, as compared to 51 percent for Republicans and 57 percent for independents).

Surveying 952 people in Sweden and the United States, William Hagman, David Anderson, Daniel Vastfjall and Gustav Tinghog similarly find that strong majorities of both Swedes and Americans support a wide variety of nudges.¹⁶ The significance of the conclusion is fortified by the fact that along many dimensions, Swedes and Americans differ – and yet their ethical evaluations are remarkably similar. Consider five examples, each involving classic nudges, in use or seriously proposed in many nations.

1. Avoiding tax evasion (appealing to conscience). *Many countries have a problem with its citizens not paying taxes, which costs society a considerable amount of money. Some countries have therefore started to send out information to the taxpayers with the encouraging message “To pay your taxes is the right thing to do”. The idea with this intervention is to give tax evaders a bad conscience and therefore increase their motivation to pay their taxes.*

2. Smoking discouragement (graphic warnings). *Smoking often leads to addiction and has a negative effect on the health of the individual. To more clearly show the negative effects of smoking, many countries have started to add deterrent pictures on the cigarette packages. These images display damaged organs that can be a consequence of long term smoking. The idea with this intervention is to discourage people to start smoking and motivate people that are smokers to quit.*

3. Cafeteria (accessibility). *Overconsumption of calorie rich food can lead to a deteriorating health. In an attempt to get their employees to eat healthier, a company rearranged its cafeteria. Healthy food was placed at eye-level and easily available for the visitors of the cafeteria. Unhealthy food, such as candy or snacks was placed behind the counter to make them less visible and accessible for the visitors in the cafeteria. The idea with this intervention is to encourage the consumption of healthier alternatives to improve the health of the employees.*

4. Energy conservation (social norms). *Most households today are over consuming energy, which leads to a waste of resources both for the household and society. Therefore energy companies that succeed in decreasing the average energy consumption among households receive government subsidies. To motivate households to lower energy consumption, an energy company attached some complementary information to the energy bill. The information added contained a comparison of energy consumption between the customer’s household and other households in the neighborhood.*

¹⁶ William Hagman et al., Public Views on Policies Involving Nudges, Review of Philosophy and Psychology (forthcoming 2015), available at <http://www.iei.liu.se/nek/forskning/jedi-lab/1.630217/Nudge20150417.pdf>.

If the customer's energy consumption was lower than the neighbors', a happy smiley face was shown on the bill. However, if the customer's energy consumption was higher than the neighbors', a sad face was shown. The idea with this intervention is that the feedback that these faces give will have a positive effect on the energy consumption of the households.

5. Food labeling (disclosure). *It can be difficult to tell which food products that are healthy and which are not, therefore a food chain started to label their products with stickers which look similar to a green and red stoplight or traffic signal. Healthy food, which is rich in minerals, vitamins, and has a low amount of fat and sugar are marked with a green tag. Unhealthy food, which is rich in fat and sugar, and has a low amount for minerals and vitamins receive a red tag. The idea with this intervention is to make it easier to make healthy choices.*

Hagman et al. find that over 80 percent of both Swedes and Americans think that the tax evasion policy is acceptable. Over eighty percent also favor disclosure to promote healthy choices (86.9 percent of Swedes, 83.8 of Americans). Hagman et al. find comparably high levels of support for both smoking discouragement policy (81 percent of Swedes, 72.6 percent of Americans) and cafeteria redesign (82.6 percent of Swedes and 76.4 percent of Americans). About two-thirds of both Swedes (66.4 percent) and Americans (67.1 percent) support the energy conservation nudge.

Consistent with expectations, Swedes are somewhat more enthusiastic than Americans about nudges, but only two of the eight tested nudges fail to attract majority support in either country, with 42.9 percent and 45.7 percent of Americans (but over 60 percent of Swedes) favoring these:

4. Organ donation (default rule). *There is currently a lack of organ donors in many countries. In some places, to become an organ donor the individual has to make an active choice and register as an organ donor with the appropriate authority. If no choice is registered, the individual is assumed to be unwilling to donate in event of an accident (so called Opt-In). In previous surveys most people report that they are willing to be an organ donor but have not registered.*

One way to increase the number of organ donors could be to automatically enroll people as organ donors unless otherwise specified (so called Opt-Out). In other words, it is up to the individual to register at the appropriate authority if they are unwilling to donate their organs in the event of an accident. The aim with this intervention (Opt-Out) is to increase the number of organ donors.

5. Climate compensation (default rule). *Carbon dioxide emissions in connection with flying have a negative effect on the environment. To compensate for this, there is usually a voluntary fee that travelers can add to the final price. The money from this fee goes to projects to reduce emissions of carbon dioxide to a corresponding level of the emission caused by the flight. To increase the number of travelers that choose to pay the climate compensation fee, it can automatically be added to the final price. Then, if a*

traveler does not want to pay the fee, the traveler instead has to make an active choice not to pay the fee (also known as Opt-Out). The idea with this intervention (Opt-Out) is to increase the number of travelers that compensate for climate.

What accounts for the majority's rejection of these nudges in the United States (and significant opposition in Sweden as well)? I speculate that the answer lies in this ethical principle: *choice architects should not use people's inertia or inattention against them*. For decisions that have a significant degree of moral sensitivity (organ donation) or sheer cost (climate change compensation), many people reject a default and favor active choosing. The apparent idea – for which more empirical testing would be desirable -- is that if a default rule would lead people to end up with an outcome that is morally troubling (to them) or expensive (for them), that rule is objectionable and active choosing is much better.¹⁷

We could confidently predict widespread disapproval of a default rule announcing that voters will, by default, be taken to support incumbent politicians (subject to opt out), or that employees will, by default, give 10 percent of their earnings to their employer's children, or to their employer's favorite charity. It is reasonable to speculate that in evaluating defaults, people are sensitive to the question whether the result is to track the desires and values of all or most of the population that is subject to them.

That lesson is a significant one, but the most important finding is the apparently widespread endorsement of nudges, whether the goal is to protect third parties (as in the case of tax evasion) or the self (as in the case of smoking discouragement). In general, Hagman et al. find that larger percentages of people are more supportive of third-party nudges, but the difference is not large, and many nudges that are designed to protect the self receive substantial support. Not surprisingly, Hagman et al. also find that those with an individualistic worldview are (somewhat) less likely to embrace nudges. More strikingly, they find that respondents *who are more prone to analytical thinking are less likely to see nudges as intruding on freedom of choice*. It may be that analytical thinkers are more able to see that the relevant nudges sufficiently preserve freedom, whereas others have a more immediate and visceral reaction, leading to a more skeptical (and erroneous) conclusion.

The Swedish-U.S. differences remain noteworthy even in the midst of the general agreement between people in the two nations. It would of course be valuable to test a wider array of nudges and to see what kinds of division might emerge. We could explore whether there are systematic differences between “harm-to-self” nudges and “harm-to-others” nudges. For reason suggested above, the former might well be more controversial than the latter, at least as a general rule. But everything is likely to depend on the nature of the particular nudge. Some “harm-to-self” nudges, such as calorie labels and automatic enrollment in pension plans, might well attract more support than some “harm-to-others” nudges, such as default rules that promote use of environmentally-friendly (but expensive) energy providers. We could also see whether people reject particular

¹⁷ For relevant discussion, see Cass R. Sunstein, *Choosing Not To Choose* (2015).

categories of nudges -- for example, those that seem to involve especially personal or intimate choices. It is fair to assume that both Swedes and Americans would be unhappy with a system of “default spouses” or “default religions” or “default political affiliations.”

III. Nudging System 1

In behavioral science, it has become standard to distinguish between two families of cognitive operations in the human mind: System 1, which is fast, automatic, and intuitive, and System 2, which is slow, calculative, and deliberative.¹⁸ System 1 is distinctly associated with identifiable behavioral biases, especially in unfamiliar or unusual situations.¹⁹ To be sure, System 2 can and does err; some math problems are hard, and people can perform exceedingly well through fast and frugal heuristics.²⁰ Professional tennis players have educated Systems 1, and they tend to know exactly what shot to hit in an instant; something similar can be said about experienced doctors, lawyers, and engineers.²¹ But there is a close connection with behavioral biases and the automatic system.

Because of the operation of System 1, for example, many people show “present bias,” focusing on the short-term and downplaying the future.²² People do not deal well with probability, in part because they use heuristics, or mental shortcuts, that sometimes lead them in unfortunate directions.²³ With respect to probability, people’s intuitions can go badly wrong, in the sense that they can produce serious mistakes, including life-altering ones.²⁴ Most people also tend to be unrealistically optimistic.²⁵

With the distinction between deliberative and automatic processing in mind, we might want to distinguish between nudges that address System 1 and nudges that address System 2. Many of the most powerful arguments against nudging appear to focus on “System 1 nudges” and to suggest that they are distinctly objectionable.²⁶ If government or the private sector is attempting to influence people by targeting or exploiting their automatic systems, or by enlisting their biases, it might seem to be engaged in

¹⁸ See Daniel Kahneman, *Thinking, Fast and Slow* (2011).

¹⁹ See Kahneman, *supra* note.

²⁰ Gerd Gigerenzer et al., *Simple Heuristics That Make Us Smart* (2000). An excellent discussion of the underlying debate is Mark Kelman, *The Heuristics Debate* (2010).

²¹ See Gary Klein, *Seeing What Others Don’t* (2015); Gary Klein, *Sources of Power* (1999).

²² For references and discussion, see Sunstein, *Why Nudge?*, *supra* note.

²³ See Kahneman, *supra* note.

²⁴ For a powerful demonstration, see Daniel Chen et al., *Decision-Making under the Gambler’s Fallacy: Evidence from Asylum Judges, Loan Officers, and Baseball Umpires* (2014), available at <http://users.nber.org/~dlchen/papers/ColdHand.pdf>

²⁵ See Tali Sharot, *The Optimistic Bias* (2011).

²⁶ This is one way to read Till Grine-Yanoff and Ralph Hertwig, *Nudge Versus Boost: How Coherent Are Policy and Theory, Mind & Machines* (2015).

manipulation, and to be treating people without respect.²⁷ It might also appear to be disparaging their agency. On this view, nudges that inform System 2 might seem far better, because they help people to reflect, or to improve their deliberative capacities.²⁸

But do people actually care about the difference? Exactly when and how? These questions can be tested, as Gidon Felsen and his colleagues have shown.²⁹ Notice, for example, the difference between two scenarios, involving nudges that are designed to increase savings.³⁰

(1) The new design works like this—with every annual salary increase you are provided information in the form of a series of icons representing tropical beaches that shows how much extra leisure you are likely to be able to afford during your retirement by investing different percentages of your increased salary; larger investments now translate into more retirement savings later. You can still choose to keep the entire salary increase instead of investing it, but the information provided results in a subconsciously-driven bias towards investment; in other words, the decision to invest is made more likely as a result of subconscious deliberation. Studies have shown that implementing this policy leads to an increase in retirement savings.

(2) The new design works like this—with every annual salary increase you are provided information in the form of a detailed table of your earnings that shows how much extra money you are likely to have during your retirement by investing different percentages of your increased salary; larger investments now translate into more retirement savings later. You can still choose to keep the entire salary increase instead of investing it, but the information provided results in a consciously-driven bias towards long-term investment; in other words, the decision to invest is made more likely as a result of conscious deliberation. Studies have shown that implementing this policy leads to an increase in retirement savings.

Or consider the difference between two approaches designed to promote healthy eating:

(1) The new design works like this — the cafeteria has been revamped so that unhealthy foods, such as candy bars, potato chips, and the like are not as conveniently located. You can still choose whichever foods you would like, but moving the location of the unhealthy food in the cafeteria results in a subconsciously-driven bias towards healthy eating choices; in other words, the decision to eat healthy foods is made more likely without the need for conscious deliberation. Studies have shown that implementing this policy leads to healthier eating habits.

²⁷ See Waldron, *supra* note.

²⁸ See Grine-Yanoff and Hertwig, *supra* note.

²⁹ Gidon Felsen et al., *Decisional Enhancement and Autonomy: Public Attitudes Toward Overt and Covert Nudges*, 8 *Judgment and Decision Making* 203 (2012).

³⁰ *Id.*

(2) The new design works like this - the cafeteria has been revamped so that all foods have their nutritional content clearly displayed. You can still choose whichever foods you would like, but the nutritional information results in a consciously-driven bias towards healthy eating choices; in other words, the decision to eat healthy foods is made more likely as a result of conscious deliberation. Studies have shown that implementing this policy leads to healthier eating habits.

Such questions allow for a test of this hypothesis, connected with the earlier discussion of manipulation: *people are more likely to object to nudges that appeal to unconscious or subconscious processes.* Surveying 2,775 people in Canada and the United States, Felsen et al. find that people do indeed show some preference for nudges that lack that characteristic.³¹ In a range of cases – involving not only healthy eating and savings but also wise investing and prudent online purchasing – people are somewhat more likely to favor approaches that involve reflection and deliberation.

In the experimental design, subjects were not asked directly whether they preferred one nudge to another (a potentially biasing frame). Instead they were asked whether they would be more or less likely to accept a job offer from a company that offered a particular nudge or a company that did not (the neutral condition). For about half of the respondents, the comparison was between a System 1 nudge and the neutral condition. For the other half, the comparison was between a System 2 nudge and the neutral condition. The relevant scale ranged from 1-10, with 1 meaning “much less” likely to accept a job offer, and 10 meaning “much more.” The authors compared the effect of the System 1 nudge and the System 2 nudge on people’s likelihood of accepting a job offer. They found that in aggregate, people showed an increased willingness to accept job offers with a System 2 nudge (on average, around 8 on the 1-10 scale) as compared to those with an System 1 nudge (on average, around 6 on the 1-10 scale).³²

It is important to see that while people were more favorably disposed to System 2 nudges, they found System 1 nudges to be a positive inducement as well, generally concluding that they would increase the likelihood that they would accept a job offer.³³ Nonetheless, System 2 nudges were preferred. Why? A possible reason is that people do not like being manipulated and think that when nudges appeal to unconscious or subconscious processes, they offend individual dignity. In strong support of this speculation, Felsen et al. find that when conscious processing is involved, people believe that the resulting decisions are more “authentic,” evidently in the sense that those decisions reflect the chooser’s own agency. They conclude that their evidence supports “the idea that preserving the individual’s capacity for making authentic decisions is an important condition for the acceptability of decisional enhancement programs.”³⁴

³¹ Id.

³² Id. at 205. I am simplifying some aspects of their analysis; see id. for details.

³³ Id. at 205.

³⁴ Id. at 206.

Recall, however, that the difference in people's reactions is modest; it is not as if people systematically approve System 2 nudges and systematically disapprove System 1 nudges. Moreover, there is reason to suspect that when people believe that some kind of behavioral bias – such as a self-control problem -- is genuinely at work, they will become more receptive to nudges that target unconscious or subconscious processes. Felsen et al. find intriguing support for this suspicion, for in one scenario (involving eating), people were equally favorable to System 1 and System 2 nudges *when they wanted help*.³⁵ (The particular question, with answers on a 9-point scale, was “To what extent do you feel like you could use help making healthier eating choices in the face of the availability of unhealthy but tasty foods?”) If people are aware that they are suffering from a problem of self-control, and if they want to overcome that problem, an approach that targets System 1 might be unobjectionable or even welcome. The conclusion might well be fortified if people believe that existing decisions are already a product of unconscious processing. In such cases, it might be acceptable to meet fire with fire.

As Felsen et al. suggest, “covertly influencing decision processes such that the resulting decision is *aligned with higher-order* desires may actually enhance autonomy, especially in situations in which the target population is known to want help with a given behavior.³⁶ They suggest that “respondents who wanted help with eating decisions may have been more likely to recognize that food choices are often subconsciously driven, and were therefore just as likely to favor the decisional enhancement program with covert influences as the program with overt influences, whereas respondents who did not want help with food choices reverted to the expected preference for overt influences.”³⁷ It would be valuable to obtain much more evidence on this question, but we might speculate that people's evaluations of System 1 nudges would very much depend on whether they believe that it is necessary to counteract a self-control problem.

III. Partisan Nudge Bias

Do political judgments matter to people's assessment of nudges?³⁸ Casual observation suggests that they do. When the Obama Administration uses behaviorally informed tools, those who are inclined to oppose the Obama Administration are not likely to love those tools.³⁹ Consider this hypothesis: *At least across a wide range,*

³⁵ Id. at 206-207. Note that when people wanted help, they were (not surprisingly) more likely to favor some kind of nudge over the neutral option. Id. at 206. But in most of the scenarios, they continued to show a relative preference for the System 2 nudge; the healthy eating scenario was the exception. Id. at 207 (“the less respondents wanted help, the more favorable they were to the conscious than the subconscious influence”).

³⁶ Id.

³⁷ Id. at 208.

³⁸ David Tannenbaum et al., *On the Misplaced Politics of Behavioral Policy Interventions* (unpublished manuscript 2014), available at <http://home.uchicago.edu/~davetannenbaum/documents/partisan%20nudge%20bias.pdf>.

³⁹ See <http://www.thenewamerican.com/usnews/constitution/item/16408-obama-behavior-team-to-nudge-u-s-toward-government-goals>

people have no considered view on nudges as such. Their evaluations turn on whether they approve of the politics of the particular nudge, or the particular nudges that come to mind. The hypothesis draws support from related evidence that on many general questions, including institutional ones, people lack clear convictions, and their judgments turn on what they think about the underlying political substance.⁴⁰

More specific evidence supports this view.⁴¹ In a series of studies, David Tannenbaum, Craig Fox, and Todd Rogers have found what they call “partisan nudge bias.” Focusing on policies favoring automatic enrollment in pension plans, they randomly assigned people to conditions in which they learned that such policies had been implemented by the Bush Administration, the Obama Administration, or an unnamed Administration. After informing participants about the policy nudge, Tannenbaum et al. reminded them that defaults could be used “across a wide range of policies beyond the illustration above” and asked how they felt, setting the particular application aside, “about actively setting default options as a general approach to public policy.”⁴²

The basic finding was that on the general question, people were much influenced by whether Bush or Obama was responsible for the particular nudge that they read about. When participants were informed that the pension default had been implemented by Obama, liberals tended to display relative support for the use of defaults as a general policy tool, whereas conservatives tended to oppose them. But when told that the same policy had been implemented by Bush, that pattern was eliminated.

Tannenbaum et al. also asked respondents about a series of nudges that had an identifiable political valence, immediately triggering disparate reactions from liberals and conservatives. These included increasing participation by low-income individuals in existing food stamp and supplemental nutrition assistance programs (liberal valence); increasing claims by high-income individuals for existing capital gains tax breaks (conservative valence); increasing participation in safe sex and effective contraception use educational programs for high-school children (liberal valence); increasing participation in intelligent design educational programs for high-school children (conservative valence); and a generic, context-free policy illustration (no valence). There were five different types of policy nudges: (1) automatic enrollment defaults, (2) implementation intentions, (3) public commitments, (4) highlighting losses, and (5) descriptive social norms. As in their first study, Tannenbaum et al. asked people about their *general* views about nudges, after seeing the relevant example. Participants were specifically reminded that the approach was general and could be used across a wide range of policies.

The result was unambiguous: People are significantly more likely to approve of nudges in general when they favor the particular political objectives used to illustrate

⁴⁰ See Eric Posner and Cass R. Sunstein, *Institutional Flip-Flops*, *Texas L Rev* (forthcoming 2015).

⁴¹ Tannenbaum et al., *supra* note.

⁴² *Id.*

them.⁴³ When the nudges were applied to traditionally liberal policies (food stamps, safe sex), liberals were relatively supportive of nudges as policy tools, while conservatives were relatively opposed to their general use.⁴⁴ This pattern reversed when those same nudges were applied to traditionally conservative policy goals (capital gains programs, intelligent design education programs).⁴⁵

Interestingly, and importantly, when nudges were attached to a *generic* policy objective, there was no association between political orientation and people's evaluation of nudges; apparently conservatives and liberals do not disagree on the general question. A particularly striking finding: While libertarians were less likely to approve of nudges than those without libertarian dispositions, attitudes about particular policies turned out to be a far more significant predictor than attitudes about libertarianism in general.

Tannenbaum et al. used the same basic strategy to test the responses of actual policymakers, consisting of U.S. city mayors and high-level public servants in state and local governments. They asked the participants to read about two kinds of automatic enrollment defaults. Half read a scenario in which low-income earners were automatically defaulted to receive supplemental food assistance benefits, and half read a scenario in which high-income earners were automatically defaulted to receive capital gains tax benefits. Policymakers were explicitly reminded that the task was the evaluation of nudges as general-purpose policy tools. The usual pattern held: The overall assessments of policymakers were greatly affected by the political valence of the examples.

In sum, “people find nudges more ethically problematic when they are applied to policy objectives they oppose, or when applied by policymakers they oppose, while they find the same nudges more acceptable when they are applied to political objectives they support or by policymakers they support.”⁴⁶ It would not of course be surprising to find that people favor nudges that support their own goals and reject nudges that undermine those goals. What is more interesting is that many people seem not to have strong or firm judgments about nudges, taken simply as such. Particular examples drive their general views -- perhaps because the examples create some kind of affective reaction to the broad category, perhaps because the examples are taken to convey information about how nudges would actually be used (which should of course bear on the overall evaluation). In this respect, people use the examples as heuristics, or mental shortcuts, in answering the broader and more difficult question. (This finding reflects a form of “attribute substitution,” of the kind that has been found in many contexts.⁴⁷)

⁴³ Id.

⁴⁴ Id.

⁴⁵ Id.

⁴⁶ Id.

⁴⁷ See Daniel Kahneman and Shane Frederick, Representativeness Revisited: Attribute Substitution in Intuitive Judgment, in *Heuristics and Biases: The Psychology of Intuitive Judgment* (Thomas Gilovich et al. eds. 2002).

There is a clear implication here for the political economy of nudging: Citizens' judgments about the ethics of nudging, and even the general enterprise, are likely to be, in significant part, an artifact of their substantive judgments about the specific directions in which they think that people are likely to be nudged. It is noteworthy that in the United Kingdom, nudging has been prominently associated with the Conservative Party (and Prime Minister David Cameron), which has likely reduced concern from the right (and perhaps heightened concern from the left).⁴⁸ To be sure, this point should not be taken too far. Even those who strongly support an incumbent president would be likely to object strenuously if he imposed a nudge that entrenched himself (as, for example, through a system of default voting). In egregious cases of manipulation, citizens of a free society (or even an unfree one) might well be outraged whatever they think of the underlying substance. But within certain limits, political assessments are likely to reflect political judgments.

IV. The Effects of Transparency About Nudging

If people are explicitly informed that they are being nudged, does their behavior change? This question does not ask about people's ethical evaluations – at least not directly. Instead it tests a seemingly plausible hypothesis, which is that if people are told that they are being nudged, they will react adversely and resist (and hence be nudged less or not at all). That hypothesis is closely connected with ethical issues: If people resist nudges when they are told about them, then we have some reason to think that they believe that nudges are ethically questionable, at least in some relevant respects.

On one view, the effectiveness of (some) nudging depends on at least a degree of nontransparency. The philosopher Luc Bovens contends broadly that the underlying psychological mechanisms “typically work better in the dark. If we tell students that the order of the food in the Cafeteria is rearranged for dietary purposes, then the intervention may be less successful. If we explain the endowment effect to employees, they may be less inclined to Save More Tomorrow.”⁴⁹ And indeed, some people have contended that nudging is ethically questionable for that reason.⁵⁰

⁴⁸ On the experience in the United Kingdom, see David Halpern, *The Nudge Unit* (forthcoming 2015).

⁴⁹ Luc Bovens, *The Ethics of Nudge*, in *Preference Change* 207 (Till Grune-Yanoff and S.O. Hansson eds. 2008).

⁵⁰ See Rebonato, *supra* note, for a series of objections, some of which involve this claim. In a related vein, Sarah Conly contends that when nudges are at work, “Rather than regarding people as generally capable of making good choices, we outmaneuver them by appealing to their irrationality, just in more fruitful ways. We concede that people can't generally make good decisions when left to their own devices, and this runs against the basic premise of liberalism, which is that we are basically rational, prudent creatures who may thus, and should thus, direct themselves autonomously.” Sarah Conly, *Against Autonomy* 30 (2011).

Even without empirical testing, we should be careful before accepting this claim. Most nudges are fully transparent, and all of them should be; they are hardly in the dark. Disclosure, reminders, warnings, uses of social norms – none of these is exactly hidden, and they need to be transparent in order to work. In general, public officials should inform people about what they are doing. But the idea of transparency is not self-defining. Is a transparency requirement satisfied merely because the nudge is not secret? Is there also an obligation to inform people that they are being nudged? What, exactly, does that mean? Is there an obligation to inform people about the specific psychological mechanisms that make nudges effective? (That appears to be Bovens' concern.) Is there any such obligation at the time that people are choosing?

We might ask, with Bovens, whether explicit warnings (“you are about to be nudged” or “we are exploiting the endowment effect”) would turn out to undermine the whole enterprise. It is reasonable to wonder about the effects of certain kinds of transparency. Consider this: “We know that people tend to do what other people do, and so we are telling you about the social norm in order to get you to do what other people do.” Would that kind of transparency prove self-defeating? Or consider this: “We know that because of inertia, default rules stick. So we are using a default rule to affect your life -- in the hope that it will stick.” A statement of this kind might well make the nudge less effective, but perhaps not; perhaps the disclosure would increase the effect, or perhaps it would be immaterial.

There is not a great deal of evidence on these questions, but an important study by George Loewenstein, Cindy Bryce, and David Haggman offers the following finding, at least in the context of end-of-life care: *When people are specifically informed that a default rule has been put in place, and that it might be otherwise, that information has little effect on what people end up doing.*⁵¹

As Loewenstein et al. designed the experiment, people were given one of these default options. (a) “I want my health care providers and agent to pursue treatments that help me to live as long as possible, even if that means I might have more pain and suffering.” (b) “I want my health care providers and agent to pursue treatments that help relieve my pain and suffering, even if that means I might not live as long.” In the experiment, one or the other of these was pre-selected as the default, but participants could change it by selecting a different alternative and confirming the change with their initials. Note that this is an exceedingly weak default, not only in the sense that it is exceptionally simple to change it, but also in the sense that the option to switch is made highly salient to participants, so that the problem of procrastination and inertia, which often makes defaults “sticky,” is greatly reduced. More familiar defaults (for example, automatic enrollment in pension or health care plans) are more likely to stick, in part

⁵¹ See George Loewenstein, Cindy Bryce, David Haggmann & Sachin Rajpal, Warning: You Are About To Be Nudged 17 (Mar. 29, 2014) (unpublished working paper) (<http://ssrn.com/abstract=2417383>) (“[I]nforming people about default interventions in advance . . . does not significantly diminish the impact of defaults on expressed preferences in advance directives.”).

because it is simple for people to ignore the question whether to depart from them, or to decide that they will consider that question at some future time.

Here is the disclosure provided by Loewenstein et al., letting people know that they have been defaulted: “The specific focus of this research is on ‘defaults’ – decisions that go into effect if people don’t take actions to do something different. Participants in this research project have been divided into two experimental groups.” Having received this information, participants were also told: “If you have been assigned to one group, the Advance Directive you complete will have answers to questions checked that will direct health care providers to help relieve pain and suffering even it means not living as long. If you want to choose different options, you will be asked to check off different option and place your initials beside the different option you select.” Participants were informed as well that “if you have been assigned to the other group, the Advance Directive you complete will have answers to questions checked that will direct health care providers to prolong your life as much as possible, even if it means you may experience greater pain and suffering.”

Notably, this information had little effect on participants’ ultimate choices. To be sure, the default itself had a significant effect; as predicted, more people chose “comfort” when it was the default. But even when people were specifically given the disclosure, signaling “the specific focus of this research” *before* they made their choices, there was no significant consequence for where they ended up, thus “suggesting that pre-informing respondents does not diminish their tendency of sticking with the default.”⁵² A possible explanation is that participants thought something like “yeah, whatever” when they read the disclosure. For some of the same reasons that default rules stick -- inertia and inattention -- disclosures of this kind might have little influence on people’s decisions. Here, then, is a general hypothesis: *Even if people are informed that they are being nudged, the effects of the nudge will usually not be reduced, either because people do not care, or because they will not expend the effort to focus on that information.*

In some contexts, the hypothesis probably will not hold. Suppose, for example, that people were told, “Over 70 percent of people engage in behavior X. We are telling you that because we want you to engage in behavior X, and because people like to do what most people do.” In that case, the disclosure of the specific motivation for the nudge might reduce its impact. Recall that Bovens is concerned with the psychological mechanisms behind nudges; he thinks that if those mechanisms are not “in the dark,” nudges will be less effective. Loewenstein et al. did not directly test this claim. It would be interesting to know if the results found by Loewenstein et al. would have been different if people had been told something like this: “Default rules often have significant effects on behavior, because of the force of inertia, and because people often think that such rules reflect what most people do.” Perhaps people would rebel if they were informed of the mechanisms that account for the effects of defaults. An understanding of those mechanisms might lead people to be on their guard.

⁵² Id. at 10. On post-informing, see id.

If so, we would want to distinguish between two kinds of nudges: those for which disclosure of the psychological mechanisms would be perceived as innocuous and those for which such disclosure might arouse suspicion. Provision of information, reminders, and warnings generally fall in the first category. For them, the psychological mechanisms are fairly obvious, and it is hard to see why anyone would be troubled by them. For default rules and uses of social norms, it is possible that disclosure of the mechanisms would produce at least a degree of concern. That question remains to be tested.

It would also be valuable to know if the setting of end-of-life care is distinctive in this respect, and if larger effects, from the disclosure in the Loewenstein et al. experiment, would be found in other contexts. The topic of end-of-life care is both complex and unpleasant to think about, and for that reason, a default might be especially likely to stick whatever the accompanying disclosures. Consider in this regard the fact that some disclosures have little impact precisely because it is effortful to process them.⁵³ In contexts that involve less effort, and clearer antecedent preferences, default rules are less likely to stick,⁵⁴ and disclosures might make them less sticky still.

Almost certainly, a great deal depends on whether participants believe that choice architects are trustworthy, and also on whether they are generally rebellious “types.” If people are told that a self-interested choice architect has chosen a default rule for them, and that default rules usually stick and were chosen for that reason, they might be willing to reject the rule in question. The idea of “reactance” points to this possibility: People do not like being controlled or coerced, and if they think that their options have been truncated, they might do whatever they can to take their own path.⁵⁵

Even for default rules, the possibility of reactance must be taken into account. But the findings by Loewenstein et al. make it reasonable to speculate that at least in many contexts, disclosure that a nudge is in place, and could be otherwise, would not much affect outcomes.

V. Five Conclusions

A great deal remains to be learned about people’s assessment of nudges, but five conclusions seem likely to hold. First, there is widespread support for nudges, at least of the kind that democratic societies have adopted or seriously considered in the recent past. Second, that support diminishes when people fear that because of inertia and inattention, citizens might end up with outcomes that they do not like. Third, there is somewhat

⁵³ Ryan Bubb, TMI? Why the Optimal Architecture of Disclosure Remains TBD, Mich L Rev (forthcoming 2015), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2483983

⁵⁴ See Cass R. Sunstein, Choosing Not To Choose (2015).

⁵⁵ See Sharon Brehm and Jack Brehm, Psychological Reactance: A Theory of Freedom and Control (1981); Louisa Pavey & Paul Sparks, *Reactance, Autonomy and Paths to Persuasion: Examining Perceptions of Threats to Freedom and Informational Value*, 33 MOTIVATION & EMOTION 277 (2009).

greater support for System 2 nudges than for System 1 nudges, even though many people will approve of the latter as well (especially if they are meant to combat self-control problems).

Fourth, people's assessment of nudges in general will be greatly affected by the political valence of the particular nudges that they have in mind (or that are brought to their minds). Fifth, transparency about nudging should not, in general, reduce the effectiveness of nudges, because most nudges are already transparent, and because people will not, in general, rebel against nudges. The principal qualification to the last point is that in some cases, reactance cannot be ruled out, especially if people do not like or trust the choice architect, and if they believe that they are being tricked or manipulated.

