

ARTICLE

The rhetoric of reaction redux

Cass R. Sunstein

Robert Walmsley University Professor, Harvard University, Cambridge, MA, USA
Email: csunstei@law.harvard.edu

(Received 19 July 2022; accepted 19 July 2022; first published online 23 August 2022)

Abstract

In *The Rhetoric of Reaction*, published in 1991, Albert Hirschman identified three standard objections to reform proposals: perversity, futility and jeopardy. In Hirschman's account, these objections define reactionary rhetoric. A proposal would be "perverse" if it would aggravate the very problem it is meant to solve; it would be "futile" if it would not even dent the problem; it would produce "jeopardy" if it would endanger some other goal or value (such as liberty or economic growth). The rhetoric of reaction comes from both left and right, though in Hirschman's account, it is a special favorite of the right. In recent years, the perversity, futility and jeopardy theses have often been invoked to challenge reforms, including nudges. While the three theses are sometimes supported by the evidence, they are often evidence-free speculations, confirming Hirschman's suggestion that the rhetoric of reaction has "a certain elementary sophistication and paradoxical quality that carry conviction for those who are in search of instant insights and utter certainties."

Keywords: perversity; jeopardy; futility; unintended consequences

Just once in a while one would like to see them a little less disabused and bitter, with perhaps a dash of that naivete they are so bent on denouncing, with some openness to the unexpected, the possible.

Albert Hirschman

Futility, perversity, jeopardy

In *The Rhetoric of Reaction*, published in 1991, Albert Hirschman sought to catalogue what he called "two hundred years of reactionary rhetoric" (p. 1). He identified three "reactionary" objections to reform proposals: perversity, futility and jeopardy (p. 7). In Hirschman's account, the three objections can be found in many times and places; they have a mechanical quality. He emphasized that perversity, futility and jeopardy theses do not come only from reactionaries. They can be invoked whenever a new proposal is offered. But he urged that "the arguments are most typical of conservative attacks on existing or proposed progressive policies" (pp. 6–7).

Of the three, the perversity thesis is the most dramatic (and clever); it suggests that “any purposive action to improve some feature of the political, social, or economic order only serves to exacerbate the condition one wishes to remedy” (p. 7). The power of the perversity thesis lies in the claim that the relevant reform “will produce, via a chain of unintended consequences, the exact contrary of the objective being proclaimed and pursued” – what Hirschman called “a daring intellectual maneuver” (p. 11). According to the perversity thesis, “the attack to push society in a certain direction will result in its moving all right, but in the opposite direction” (p. 11). If “everything backfires,” then (for example) “social welfare programs will create more, rather than less, poverty” (because people will not have an incentive to work), and efforts to require people to wear seatbelts will cost lives (because people will drive more dangerously) (pp. 12 and 41).¹ Hirschman shows that the perversity argument played a central role in reactionary responses to the French Revolution; in attacks on universal suffrage; and in objections to the modern welfare state, above all in the field of economics (pp. 12–35).

In drawing attention to the pervasiveness of the perversity claim, Hirschman did not mean to insist that the claim was always wrong. But he did insist that “it is unlikely to exist ‘out there’ to anything like the extent that it is claimed” (p. 35). On this count, Hirschman wrote bluntly: Could reactionaries “be embracing the perverse effect for the express purpose of feeling good about themselves? Are they not being unduly arrogant when they are portraying ordinary humans as groping in the dark, while in contrast they themselves are made to look so remarkably perspicacious?” (p. 36). He worried that “the effect is likely to be invoked for reasons that have little to do with its intrinsic truth value,” in part because “there are many unintended consequences that are *welcome* rather than the opposite” (p. 38). He insisted that the perversity thesis “has a certain elementary sophistication and paradoxical quality that carry conviction for those who are in search of instant insights and utter certainties” (p. 43).

The futility thesis is the simplest and perhaps the most intuitive; it suggests “that attempts at social transformation will be unavailing, that they will simply fail to ‘make a dent’” (p. 7). For those who invoke the futility thesis, any change will be “surface, façade, cosmetic, hence illusory, as the ‘deep’ structures of society remain wholly untouched” (p. 42). Hirschman illustrates the futility thesis with a joke from Eastern Europe after the installation of Communist government in the aftermath of World War II: “What is the difference between capitalism and socialism?” The answer: “In capitalism, man exploits man; in communism, it’s the other way round” (p. 44).

Tocqueville’s elaborate work on the French Revolution can be seen as a case study in the futility thesis, suggesting as it did that the magnitude of a seemingly epochal change was relatively modest (Hirschman, 1991, pp. 45–50). Many attacks on universal suffrage have urged that it would be a cosmetic reform and achieve little or nothing (pp. 50–60). So too, reactionaries have objected that the real beneficiaries of the

¹The offsetting effect of more dangerous driving is often described as “the Peltzman Effect” (Peltzman, 1975); Peltzman’s own finding is one of futility, not perversity. In the same vein, see Graham and Wiener (1997). For a Hirschman-like response, see Riscoff and Revesz (2002).

welfare state are not the poor, but “other social groups with more clout,” including “a large group of administrators, social workers, and sundry professionals ... wholly dedicated to expanding their bureaus and perquisites” (pp. 60–69). In all cases, the objection is that “the basic structures of society” are unaltered (p. 72). Hirschman has something direct to say to those drawn to the futility thesis (he could have said it to those drawn to the perversity thesis and the jeopardy thesis as well): “Just once in a while one would like to see them a little less disabused and bitter, with perhaps a dash of that naivete they are so bent on denouncing, with some openness to the unexpected, the possible” (p. 80).

The jeopardy thesis is the least straightforward; it suggests that “the cost of the proposed change or reform is too high as it endangered some previous, precious accomplishment” (p. 7). The basic idea is that while the proposed change might be desirable in itself, it “involves unacceptable cost or consequences of one sort or another” (p. 81). The Principle of the Dangerous Wedge is an example; it is more familiarly known as a “slippery slope” argument, which means that if you do X (when X is admittedly desirable), you will also end up doing Y (when Y is truly terrible) (Hirschman, 1991, p. 82 citing Cornford, 1908). On one view, for example, a small step toward regulating gun ownership should be resisted, because it will lead to the confiscation of guns. On another view, democracy threatens liberty (as, for example, when majorities want to punish dissenters). On another view, the welfare state threatens liberty (as, for example, when it results in violations of freedom of contract). On another view, modern regulatory agencies threaten democracy (because they give power to unelected bureaucrats). Or consider these words from Gustave Le Bon:

Had democracies possessed the power they wield to-day at the time of the invention of mechanical looms or of the introduction of steam-power and of railways, the realisation of these inventions would have been impossible, or would have been achieved at the cost of revolutions and repeated massacres. It is fortunate for the progress of civilisation that the power of crowds only began to exist when the great discoveries of science and industry had already been effected (2002, p. 35).

Friedrich Hayek often invoked the jeopardy thesis: “Freedom is critically threatened when the government is given exclusive power to provide certain services – power which, in order to achieve its purposes, it must use for the discretionary coercion of individuals” (Hirschman, 1991, p. 114 quoting Hayek, 1939). Hirschman did not claim that the jeopardy thesis was always false, but he responded to it with a combination of bemusement and rage. Because the thesis is offered much more often than it turns out to be true, “there must be some inherent intellectual attraction in advancing” it (p. 164). Sometimes a reform does not, in fact, jeopardize some other reform. Sometimes it creates a kind of complementarity or mutual support. For example, the Earned Income Tax Credit might supplement other antipoverty programs and not compromise them at all. Graphic warnings on cigarette packages might supplement and even spur cigarette taxes, and not displace them. Democracy might turn out to promote innovation, rather than reducing it. A single reform, said to jeopardize

some other practice or option, might actually make it more likely or more real, or larger still.

Hirschman almost titled his book “The Rhetoric of Intransigence,” and while that would have been an unlovely title, it was his central preoccupation, and his final chapter is called, simply, “Beyond Intransigence.” In that chapter, Hirschman takes off the gloves. He suggests that with his procedure in mind, “some ‘deep thinkers’ who had invariably presented their ideas as original and brilliant insights are made to look rather less impressive, and sometimes even comical” (p. 164). He notes that his “more basic intent” was “to establish some presumption, through the demonstration of repetition in basic argument, that the standard ‘reactionary’ reasoning, as here exhibited, is frequently *faulty*” (pp. 165–166). One reason is that the arguments “are invoked time and again almost routinely to cover a wide variety of real situations” (p. 166). Hirschman’s ultimate goal “is to move public discourse beyond extreme, intransigent postures,” in “the hope that in the process our debates will become more ‘democracy friendly’” (p. 168).

Hirschman’s book appeared over 30 years ago, but it is directly relevant to countless current debates. In some ways, ours is the Era of Reactionary Rhetoric, in precisely Hirschman’s sense. To reduce poverty, for example, a nation might significantly increase the minimum wage, but (it is said) the effect of the reduction would be to discourage employers from hiring people at the bottom of the economic ladder; the effect would be perverse. To reduce gun violence, a nation might require background checks for gun purchases, but (it is said) the requirement would do no good at all; it would be futile. To combat the risks associated with COVID-19, a nation might impose a mask mandate, but (it is said) that step would jeopardize individual freedom. Debates over greenhouse gas reductions are characterized by the futility thesis (*if just one nation reduces emissions, it will do no good!*) and the jeopardy thesis (*if we limit greenhouse gas emissions, we will punish poor countries, and poor people, the most!*). We can find the perversity, futility and jeopardy arguments in countless debates – and while they might come most frequently from the right, they are common on the left as well (as in the claim that nominally progressive actions, such as increases in funding for social welfare programs, will do little or nothing to combat structural poverty or structural racism).

From the standpoint of behavioral science, the rhetoric of reaction might usefully be explored in terms of motivated reasoning or more particularly “desirability bias” (see Tappin *et al.*, 2018). The basic idea is that people are far more likely to believe things to be true if they *want* those things to be true. For some people, it is delightful to think that some reform will be futile or perverse, and it is upsetting or destabilizing to think that the same reform will be highly effective. A claim of futility or perversity is highly welcome. We can easily read Hirschman to have drawn attention to the motivated character of the rhetoric of reaction, and to have suggested that motivations should be put to one side in favor of investigation of actual evidence.

Nudges and nudging, very briefly

Nudges are interventions that steer people in particular directions but that also allow them to go their own way (see Thaler & Sunstein, 2021). A reminder is a nudge; so is

a warning. A GPS nudges; a default rule nudges. Some nudges are educative; consider labels and warnings. Other nudges are architectural; consider automatic enrollment or website design that places certain options first or in large font. To qualify as a nudge, an intervention must not impose significant material incentives. A subsidy is not a nudge; a tax is not a nudge; a fine or a jail sentence is not a nudge. To count as such, a nudge must fully preserve freedom of choice. If an intervention imposes significant material costs on choosers, it might of course be justified, but it is not a nudge. Some nudges work because they inform people; other nudges work because they make certain choices easier; still other nudges work because of the power of inertia and procrastination.

There has been a great deal of work on the actual effects of nudges (see, e.g., Arno & Thomas, 2016; ; DellaVinga & Linos, 2022; Mertens *et al.*, 2022). In some cases, they have a massive impact; changes in the default rule, from opt-in to opt-out, sometimes fall in that category (see Momsen & Stoerk, 2014, p. 380; Jachimowicz *et al.*, 2019, p. 164; Sarin, 2019, p. 1560; Bergman *et al.*, forthcoming 2022). In some cases, nudges have a modest but real impact; disclosure of information (see Ellison *et al.*, 2013), reminders, and changes in “framing” sometimes have statistically significant effects, but they are hardly massive (see Choudhary *et al.*, 2019, p. 16; DellaVinga & Linos, 2022, p. 92). In some cases, nudges have no measurable effect at all (Elbel *et al.*, 2009; Kristal & Whillans, 2020, p. 171; Dewies *et al.*, 2021, p. 221; Rela, 2022). In some cases, nudges have “boomerang effects”; they produce the opposite of their intended result (see, e.g., Martineau *et al.*, 2011, p. 264; Bolton *et al.*, 2020, pp. 3–4; Osman *et al.*, 2020).² In some cases, nudges have the intended effect, but they also impose costs, because (for example) people suffer welfare losses as a result of the intervention (see Allcott & Kessler, 2019; Thunstrom, 2019). – the “ruined popcorn” effect, as we might call it (see Sunstein, 2019). The welfare measure is the right one (see Sunstein, forthcoming 2023), and while we know less than we should about the welfare effects of the full universe of nudges, many of them almost certainly increase welfare.³

At the present time, many of the various findings offer a positive picture of the effects of existing nudges (see Benartzi *et al.*, 2017; Allcott & Kessler, 2019; Thunstrom, 2019; DellaVinga & Linos, 2022, p. 90), while also suggesting vindication, on some occasions, of the perversity thesis, the futility thesis and the jeopardy thesis. Some nudges do have little or no impact; some nudges do compromise other values; a few nudges do aggravate the very problem they are meant to solve. It is consistent with Hirschman’s plea for “democracy friendly” debates to suggest that we need

²On minimum payment requirements and credit card payments, see Stewart (2009), p. 39; Wang and Keyes (2014), pp. 4–5.

³A clean and simple example is Interim Rule: Direct Certification and Certification of Homeless, Migrant and Runaway Children for Free School Meals, 2011; this is the conclusion of Allcott and Kessler (2019) and Thunstrom (2019). For a study suggesting that “claw-back” nudges cause net increases in welfare, see Bulte *et al.* (2020), pp. 2348 and 2350. An important qualification involves the deliberate uses of behavioral science, or behavioral economics, to take advantage of consumers, employers, investors, or others. “Dark patterns” often involve such uses (see Luguri & Strahilevitz, 2021). Uses of behavioral science to take advantage of (for example) limited attention or present bias raise important questions, but unless we do some fancy footwork, we are not really speaking of futility, perversity or jeopardy. On the normative questions, (see Luguri & Strahilevitz, 2021; Sunstein, 2021).

much more in the way of empirical testing. In fact we might offer a (friendly) amendment to Hirschman's own rhetoric. A great advantage of empirical testing is that we end up with answers and inquiries, not debates. We can identify testable hypotheses and then test them. We might ultimately end up with a kind of "map" to various effects, with an understanding of what we are likely to find, and when we are likely to find it (see Cadario & Chandon, 2020, p. 459). If so, we might not "debate" at all, and our debates, if they occur, might well be relatively low-key. They would not involve intransigence. On the contrary, they would be characterized by openness and learning.

Of tweaks, slippery slopes and distractions

Some discussions of nudges and nudging are not low-key. In some cases, the perversity thesis, the futility thesis and the jeopardy thesis have been in something like screaming bright lights.

As applied to nudges, the futility thesis consists of the claim that nudges are mere "tweaks," and have little or no effect on social problems (Ben-Shahar & Schneider, 2014). On one view, they should be bracketed or abandoned for that reason (see Ben-Shahar & Schneider, 2014; *Nature Human Behavior* Editorial, 2020, p. 121). If, for example, the goal is to improve financial choices, disclosure requirements might do little or nothing (see Ben-Shahar & Schneider, 2014). If the goal is to increase savings in retirement, automatic enrollment in pension plans might do far too little; a savings mandate might be much better (see Bubb & Pildes, 2014, p. 1609). If the goal is to reduce obesity, calorie labels might be thin gruel (so to speak); they might have no effect at all (see Kiszko *et al.*, 2014, pp. 1252–1255). If the goal is to combat climate change (it might be said, and has been said), the focus should be on carbon taxes, subsidies, and regulatory mandates, not on home energy reports and energy-efficiency labels.⁴

The jeopardy thesis takes several forms. On one view, the choice architects who try to come up with good nudges lack the necessary information; they face the same "knowledge problem" as do planners of all kinds (see Glaeser, 2006, p. 144). For that reason, they are likely to blunder, in the sense that they will nudge people in the wrong directions (see Rebonato, 2012). Nudges are often a form of paternalism, and likely to be ill-considered for that reason (Rizzo & Whitman, 2020). Individuals care about a large number of things, not only (for example) health and safety, and if outsiders nudge people to be healthier and safer, they will compromise what individuals actually care about.⁵ On another view, nudges are or can be a form of manipulation, and in that sense, they might compromise autonomy. They might formally protect freedom of choice, but they might also operate as *de facto* mandates (Bubb & Pildes, 2014, p. 1599).

The perversity thesis is the most distinctive. For example, Ryan Bubb and Richard Pildes have argued that automatic enrollment in 401(k) plans has often turned out to be perverse, in the sense that in the aggregate, it leads to a situation in which savings go down rather than up (2014, p. 1609). The reason is that even if automatic enrollment leads to higher participation rates, a low default contribution rate (3 percent) can lead people to save less than they would if they made choices on their own –

⁴On some of the ethical issues here, see Siipi and Koi (2022), pp. 234–235.

⁵See Rebonato (2012), p. 144 which remains the best treatment.

which means that a significant increase in participation rates can lead to a significant reduction in total savings. Simply put: More nudging, less savings. Pointing to the risk of both futility and perversity, Bubb and Pildes argue in favor of mandates rather than nudges; in their view, mandates are more likely to be successful, and less likely to have boomerang effects.⁶

A generalization of the perversity argument comes from Nick Chater and George Loewenstein, who claim that focusing on individual behavior is something like a catastrophic error (2022, pp. 3–5). Chater and Loewenstein believe that it is important to distinguish between “i-frames,” focused on individuals, and “s-frames,” focused on “the system of rules, norms and institutions by which we live” (p. 3).⁷ In their account, nudges are typically i-frame interventions (see p. 9). (Importantly, nudges often do not focus on individual behavior at all; consider the greenhouse gas inventory, which is designed to nudge companies toward reducing their emissions, or “scorecard diplomacy,” which is designed to nudge nations. But I bracket that point here.) Chater and Loewenstein urge that i-frame interventions “don’t fundamentally change the rules of the game, but make often subtle adjustments that promise to help cognitively frail individuals play the game better” (p. 3). They deploy both the futility thesis and the perversity thesis, with an emphasis on the latter (pp. 10–26). In their view, the effects of i-frame interventions are “often modest,” with an impact that is “slight, at best,” which means that we have “disappointing results,” and hence a form of futility (p. 3)⁸.

In a passage that could have come directly from *The Rhetoric of Reaction*, Chater and Loewenstein argue that i-frame interventions “may be counter-productive by drawing attention and support away from the kinds of systemic changes required to address most public policy problems” (p. 5).⁹ In their view, “the entire enterprise of behaviorally-inspired public policy interventions is misdirected – and has been inadvertently aligning with the objectives of opponents of s-frame reform” (p. 6). As they put it, “promoting such interventions will either fail and/or weaken support for the system-level interventions that threaten [corporate] interests” (p. 6). A strong clue (they think) is that “across a wide range of policy domains, corporate campaigns are persistently backing i-frame approaches, presumably believing that they are failing and will continue to fail” (p. 6).¹⁰

Thus “many behaviorally oriented academics with an interest in public policy, including ourselves, have inadvertently reinforced the ineffective i-frame perspective,” and “behavioral i-frame interventions have generally had disappointing results, but more importantly can have perverse unintended consequences, such as reducing

⁶A related argument can be found in Gigerenzer (2015, p. 377), to the effect that nudges, designed to increase screening for breast cancer, can lead people in bad directions. The best response to objections of this kind, when they are convincing, is to nudge better (see Sunstein, 2017).

⁷There are evident puzzles here. Consider a sugar tax, a calorie label or safety ratings for tires; are those i-frame or s-frame interventions? What if their main goal is the change the behavior of large companies?

⁸For counterevidence, see note 3.

⁹As noted in text, some nudges are, of course, s-frame interventions; consider the idea of a greenhouse gas inventory. I do not elaborate that point here.

¹⁰Is this claim true? The answer depends on the meaning of the word “persistently.” Corporations have been opposed to, or unenthusiastic about, numerous nudges. In the United States, for example, they have attacked graphic warnings for tobacco products and front-of-the-pack labels.

support for policies solutions that are well known to be effective, and leading individuals to blame themselves for problems with systemic origins” (p. 27) (Here is the rhetoric of reaction in action, with the futility thesis followed “more importantly” by perversity thesis.) Chater and Loewenstein urge that behavioral scientists should shift to “s-frame” interventions by noting people’s insensitivity to conflicts of interest; emphasizing the power of framing; and stressing people’s power of adaptation (for example, to adapt to, and to accept, s-frame reforms that they might have resisted) (pp. 26–67).

In their central argument, Chater and Loewenstein offer a noteworthy variation on what Hirschman had in mind; call it an *argument from distraction*. The central idea, which can be applied to a host of reform proposals, is that *if one focuses on ineffective or modest Reform A, one will divert time and attention from transformative Reform B, which means that one will end up, sadly and even tragically, with Reform A when one would have otherwise ended up with Reform B*. The argument from distraction deserves an honored place in the rhetoric of reaction. Needless to say, it can be offered by the left or the right. If a reformer proposes a subsidy to promote purchases of electric vehicles, the proposal might crowd out discussion of a ban on gasoline-powered vehicles; if a reformer proposes cigarette warnings, the proposal might crowd out discussion of a substantial increase in cigarette taxes; if a reformer proposes automatic enrollment in retirement plans, the proposal might crowd out discussion of mandatory savings.¹¹

In principle, of course, the argument from distraction might be true, but it might also be false. It might even point in exactly the wrong direction. Support for Reform A might be a first step toward, and make possible, Reform B (a possibility to which I will return).

Testing

Hirschman was mostly concerned with rhetoric, not with substance. But the rhetoric of reaction raises substantive issues. What is the best way to deal with the three theses? The simplest answer is to test them empirically. What does the evidence actually tell us?

The futility thesis is the most straightforward. Do the relevant interventions have no effect? In fact? It is noteworthy that in discussions of many interventions, including nudges, the idea of futility does not generally refer to *no* effect. More often, it

¹¹The evidentiary support for this kind of claim is extraordinarily weak. To say the least, the findings in Hagmann *et al.* (2019), p. 485 figure 1 do not tell us that any nation, or any community, has not adopted a carbon tax because of the availability of a nudge. In a survey, people might not favor Aggressive Reform A if they are also given the option of Less Aggressive Reform B. But does that tell us that the United States, Ireland, China, Mexico, or Italy rejected (say) an aggressive fuel economy mandate *because* someone proposed a less aggressive fuel economy mandate, or a public education campaign to encourage the use of public transportation? (Short answer: No.) To be sure, it is *possible* that the availability of a nudge, or an i-frame reform, has discouraged the use of a mandate, or an s-frame reform, just as it is *possible* that the availability of a civil remedy has discouraged creation of a criminal remedy, or that the availability of a 70 part per billion ceiling on levels of ozone in the ambient air has discouraged the imposition of a 65 part per billion ceiling (or for that matter a 75 part per billion ceiling). But we lack evidence that if nudges were not in the picture, or if i-frame interventions were not in the picture, nations would be adopting much more aggressive approaches to (say) climate change, air pollution, alcoholism, reduction of poverty, obesity, cigarette smoking, and animal welfare. Compare the idea that nudges “pave the way for” or “clear the ground toward” something like a ban or a mandate; that too may or may not be true.

refers to effects that seem modest and perhaps disappointingly modest, especially in view of the sheer magnitude of the problem. For example, graphic warnings might prevent a very small number of the expected annual cancer deaths from smoking; fuel economy labels might do very little to address the problem of climate change; disclosure requirements might not do much to prevent credit card companies from imposing high late fees, overuse fees and interest rate charges; increasing take-up rates for the Earned Income Tax Credit might have no significant effect on the overall poverty rate. It is often possible to characterize an effect as modest, and an initiative as futile, if the benefit is very small compared to the magnitude of the underlying problem.

Beware of that characterization. Suppose, for example, that an antismoking initiative would prevent 1000 premature deaths annually. Such an initiative would have an exceedingly modest impact on the overall problem of premature deaths from smoking (estimated to be in the vicinity of 480,000 each year in the United States alone, and seven million globally), but an initiative is not exactly futile if it would prevent 1000 premature deaths annually. The monetized benefit of that savings (assuming a value of a statistical life of \$10 million) is in the vicinity of \$10 billion; producing \$10 billion in benefits is not exactly futile. Or suppose that an initiative eliminated 100 million tons of carbon dioxide – a seemingly modest gain. If the social cost of carbon is set at \$50, we are speaking of \$5 billion in monetized benefits. A nudge that achieves \$5 billion in monetized benefits is not futile. (Nothing I have said here is meant to deny that some nudges are futile. Some are; so are some subsidies and penalties.)

If a nudge ends up aggravating the very problem it is meant to solve, it is indeed perverse, and as noted, some nudges fall in that category (see Hall & Madsen, 2022). As we have seen, however, perversity in that sense is rare. To understand the perversity thesis, a key question is at what level of abstraction to describe the goal of the nudge.¹² If the goal is to improve social welfare, a nudge might be perverse even if it works to reduce caloric intake; people might be thinner but sadder. An increase in participation rates might seem to attest to successful nudging, but if participation makes people worse off, the effect will be perverse. As noted, the welfare question is central; we need to investigate it. That is true not only of nudges but also of subsidies, taxes, mandates and bans. Nudges designed to decrease alcohol consumption might be a good idea, on welfare grounds; a ban on alcohol consumption might be a bad idea, on welfare grounds.

What about i-frames and s-frames? What about the argument from distraction? If companies support Reform A and fight fiercely against Reform B, perhaps we can infer that Reform A is ineffective and that companies support it for that reason. But it is important to be careful here. Reform A might be good, all things considered, and Reform B might be bad, all things considered; companies might be right to support the first and to resist the second. (It is hazardous to think that a reform is good because companies dislike it, or that a reform is bad because companies like it, or think that it is in their self-interest to support it.) We cannot know without investigating the details. As Hirschman saw, the perversity effect might be a fairy tale. Many unintended consequences are welcome rather than the opposite, and reforms might

¹²The perversity thesis and the jeopardy thesis tend to collapse, if we describe the relevant goals at a high level of abstraction; if we do that, jeopardy starts to look like perversity.

be complementary or mutually supportive. We can imagine the following possibilities: (1) an i-frame intervention has no effect on the adoption of s-frame interventions; (2) an i-frame intervention makes it *more* likely that policymakers will make s-frame interventions; and (3) an i-frame intervention makes it less likely that policymakers will make s-frame interventions. Chater and Loewenstein are focused on (3), but they do not have anything like good evidence that it is a frequent occurrence (2022).¹³ (Their admittedly interesting tales about the behavior of some interest groups – and speculations about issues of causation in the complex domain of policy outcomes – do not count as good evidence.)

We might go further. Many people think some version of (2): In their view, a nudge might be desirable in itself, but it will pave the way toward a mandate or a ban – a slippery slope argument. For some people, that is a problem; for others, it is something to applaud. Graphic warnings on cigarettes might, in fact, strengthen support for, or provide an opening for, increases in cigarette taxes; fuel economy labels might, in fact, strengthen support for, or pave the ground for, fuel economy requirements. As we have seen, some people emphasize a version of (3), pointing to the problem of distraction. To know whether that problem is real, we would like to have a randomized controlled trial, where the control condition involved a nudge (or i-frame) and then a set of outcomes, and the treatment condition involved no nudge (or i-frame) and then a set of outcomes. The question would be whether, in the treatment condition, we observe a set of outcomes that the nudge somehow discouraged or blocked – for example, an aggressive fuel economy rule, a carbon tax, a tax on sugar-sweetened beverages, a substantial increase in the Earned Income Tax Credit, a flat prohibition on cigarettes, a more progressive income tax, criminalization of sexual harassment, a ban on late fees and overuse fees.

Because history is run only once, there is no such randomized controlled trial. But it is worth pausing over the logic of the argument. Is it reasonable to suggest that if nudges were not in the picture, the United States, Canada and Italy would have more aggressive fuel economy rules? Would the Trump administration have favored such rules if fuel economy labels were not available? (In many nations, fuel economy nudges are complementary to aggressive fuel economy rules, and supported by exactly the same people; those who seek i-frame interventions also, and simultaneously, seek s-frame interventions. In the United States, far-reaching climate legislation, enacted in 2022, was not deterred by climate-related nudges, strongly favored by many of those who worked hard for that legislation.) Is it even plausible to suggest that if graphic warnings were not in the picture (so to speak), nations would ban cigarettes? Did New Zealand take steps toward banning cigarettes because it did not consider nudges instead? Is it plausible to suggest that if disclosure requirements were not on the table, nations would forbid credit card companies from imposing late fees and overuse fees? (The United States has heavily regulated such fees.) Is it reasonable to suggest that without fuel economy labels, the United States would have adopted a carbon tax?

The general point is that the number of factors that lie behind a decision to impose, or not to impose, a transformative reform of some kind (or what is not the same thing, to adopt an s-frame) is very large. It would be a rare case indeed in which the availability of a nudge (or an i-frame intervention) was either necessary or sufficient to block a more transformative (or s-frame) reform. It might be more plausible to suggest that the

use of a nudge (or an i-frame intervention) helped spur support for a more transformative or (s-frame) reform. Recall Hirschman's worry that the rhetoric of reaction is often invoked when it has little "truth value," and that "there are many unintended consequences that are *welcome* rather than the opposite" (1991, p. 38).

By-the-numbers

Hirschman was concerned with the by-the-numbers quality of reactionary rhetoric. As Hirschman had it, reactionaries are a little like the prisoners in the old joke:

A man goes to prison and the first night while he's laying in bed contemplating his situation, he hears someone yell out, "44!" Followed by laughter from the other prisoners. He thought that was pretty odd, then he heard someone else yell out, "72!" Followed by even more laughter.

"What's going on?" he asked his cellmate.

Well, we've all heard every joke so many times, we've given them each a number to make it easier.

"Oh," he says, "can I try?"

Sure, go ahead.

So, he yells out "102!" and the place goes nuts. People are whooping and laughing in a hysteria. He looks at his cellmate rolling on the ground with tears in his eyes from laughing so hard.

Wow, good joke huh?

Yeah! We ain't never heard that one before!

Hirschman wanted to demonstrate that all of us have heard the ones about perversity, futility and jeopardy many times before. Notwithstanding that fact, Hirschman insisted that reactionary rhetoric often produces a vigorous nod of the head – sometimes chastened, sometimes enthusiastic – even when it lacks the slightest empirical foundation. Hirschman's account also suggests that reactionary rhetoric can have an identifiable psychological source: an improbably triumphant combination of bitterness and despair.

From the right, reactionary rhetoric about nudges is marked by a cheerful insistence that all really is for the best in this best of all possible worlds. From the left, reactionary rhetoric about nudges offers something distinctive: an insistence, proud but evidence-free, that choice-preserving approaches are delaying the revolution. And these points are not, of course, limited to nudges; they can be found in objections, past and present, to reforms of many diverse kinds.

Acknowledgments. I am grateful to David Halpern and A.O. Oliver for terrific comments and to Lucia Reisch and Richard Thaler for valuable discussions. Special thanks to Julia Hammond and Heather Pincus for outstanding research assistance.

Competing interest. The author declares none.

References

Interim Rule (2011) Direct Certification and Certification of Homeless, Migrant and Runaway Children for Free School Meals. <https://www.fns.usda.gov/cn/fr-042511>.

- Allcott, H. and J. B. Kessler (2019), 'The welfare effects of nudges', *American Economic Journal: Applied Economics*, **11**(1): 236–276.
- Arno, A. and S. Thomas (2016), 'The efficacy of nudge theory strategies in influencing adult dietary behavior: a systematic review and meta-analysis', *BMC Public Health*, **16**, 676.
- Ben-Shahar, O. and C. E. Schneider (2014), *More Than You Wanted to Know: The Failure of Mandated Disclosure*. Princeton: Princeton University Press.
- Benartzi, S., J. Beshears, K. L. Milkman, C. R. Sunstein, R. H. Thaler, M. Shankar, W. Tucker-Ray, W. J. Congdon and S. Galing (2017), 'Should governments invest more in nudging?', *Psychological Science*, **28**(8): 1041–1055.
- Bergman, P., J. Lasky-Fink and T. Rogers (Forthcoming 2022), 'Simplification and defaults affect adoption and impact of technology, but decision makers do not realize it', *Organizational Behavior and Human Decision Processes*, **171**.
- Bolton, G. E., E. Dimant and U. Schmidt (2020), 'When a Nudge Backfires: Combining (im)plausible Deniability with Social and Economic Incentives to Promote Behavioral Change', CESifo Working Papers, No. 8070.
- Bubb, R. and R. H. Pildes (2014), 'How behavioral economic trims its sails and why', *Harvard Law Review*, **127**(6): 1594–1678.
- Bulte, E., J. A. List and D. van Soest (2020), 'Toward an understanding of the welfare effects of nudges: evidence from a field experiment in the workplace', *The Economic Journal*, **130**(632): 2329–2353.
- Cadario, R. and P. Chandon (2020), 'Which healthy eating nudges work best? A meta-analysis of field experiments', *Marketing Science*, **39**(3): 465–486.
- Chater, N. and G. F. Loewenstein (2022), 'The i-frame and the s-frame: how focusing on individual-level solutions has led behavioral public policy astray', *SSRN Electronic Journal*, doi:10.2139/ssrn.4046264.
- Choudhary, V., M. Shunko, S. Nutessine and S. Koo (2019), 'Nudging Drivers to Safety: Evidence from a Field Experiment', INSEAD Working Papers, No. 2020/28/TOM. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3491302.
- Cornford, F. M. (1908), *Microcosmographia Academia*. Cambridge: Bowes and Bowes.
- DellaVinga, S. and E. Linos (2022), 'RCTs to scale: comprehensive evidence from two nudge units', *Econometrica*, **90**(1): 81–116.
- Dewies, M., A. Schop-Etman, K. I. M. Rohde and S. Denktaş (2021), 'Nudging is ineffective when attitudes are unresponsive: an example from a natural field experiment', *Basic Applied Social Psychology*, **43**(4): 221–225.
- Elbel, B., R. Kersh, V. L. Brescoll and L. B. Dixon (2009), 'Calorie labeling and food choices: a first look at the effects on low-income people in New York City', *Health Affairs*, **28**(Supplement 1), w1110–21.
- Ellison, B., J. L. Lusk and D. Davis (2013), 'Looking at the label and beyond: the effects of calorie labels, health consciousness, and demographics on caloric intake in restaurant', *International Journal of Behavioral Nutrition and Physical Activity*, **10**, 21.
- Gigerenzer, G. (2015), 'On the supposed evidence for libertarian paternalism', *Review of Philosophy and Psychology*, **6**: 361–383.
- Glaeser, E. L. (2006), 'Paternalism and psychology', *The University of Chicago Law Review*, **73**(1): 133–156.
- Graham, J. D. and J. Wiener (1997), *Risk vs. Risk: Tradeoffs in Protecting Health and the Environment*. Cambridge: Harvard University Press.
- Hagmann, D., E. H. Ho and G. Loewenstein (2019), 'Nudging out support for a carbon tax', *Nature Climate Change*, **9**: 484–489.
- Hall, J. D. and J. M. Madsen (2022), 'Can behavioral interventions be too salient?', *Science*, **376**(6591), eabm3427.
- Hayek, F. A. (1939), *Freedom and the Economic System*. Public Policy Pamphlet No. 29. Chicago: University of Chicago Press.
- Hirschman, A. O. (1991), *The Rhetoric of Reaction: Perversity, Futility, Jeopardy*. Cambridge: Belknap Press of Harvard University Press.
- Jachimowicz, J. M., S. Duncan, E. U. Weber and E. J. Johnson (2019), 'When and why defaults influence decisions: a meta-analysis of default effects', *Behavioural Public Policy*, **3**(2): 159–186.
- Kiszko, K. M., O. D. Martinez, C. Abrams and B. Elbel (2014), 'The influence of calorie labeling on food orders and consumption: a review of the literature', *Journal of Community Health*, **39**(6): 1248–1269.

- Kristal, A. S. and A. V. Whillans (2020), 'What we can learn from five naturalistic field experiments that failed to shift commuter behaviour', *Nature Human Behavior*, 4: 169–176.
- Le Bon, G. (2002), *The Crowd: A Study of the Popular Mind*. Mineola, New York: Dover.
- Luguri, J. and L. Strahilevitz (2021), 'Shining a light on dark patterns', *Journal of Legal Analysis*, 13(1): 43–109.
- Martineau, T. M., D. Ogilvie, M. Roland, M. Suhrcke and M. P. Kelly (2011), 'Judging nudging: can nudging improve population health?', *BMJ*, 342: 263–265.
- Mertens, S., M. Herberz, U. J. J. Hahnel and T. Brosch (2022), 'The effectiveness of nudging: a meta-analysis of choice architecture interventions across behavioral domains', *Proceedings of the National Academy of Sciences*, 119(1), e2107346118.
- Momsen, K. and T. Stoerk (2014), 'From intention to action: can nudges help consumers to choose renewable energy?', *Energy Policy*, 74(C): 376–382.
- Nature Human Behavior Editorial (2020), 'Nudges that don't nudge', *Nature Human Behavior*, 4(121): 121.
- Osman, M., S. McLachlan, N. Fenton, M. Neil, R. Löfstedt and B. Meder (2020), 'Learning from behavioural changes that fail', *Trends in Cognitive Sciences*, 24(12): 969–980.
- Peltzman, S. (1975), 'The effects of automobile safety regulation', *Journal of Political Economy*, 83(4): 677–726.
- Rebonato, R. (2012), *Taking Liberties: A Critical Examination of Libertarian Paternalism*. New York: Palgrave Macmillan.
- Rela, N. (2022), 'On nudges that fail', *Behavioural Public Policy*, 1–15.
- Rischoff, S. and R. Revesz (2002), 'The biases of risk tradeoff analysis', *The University of Chicago Law Review*, 69(4): 1763–1836.
- Rizzo, M. J. and G. Whitman (2020), *Escaping Paternalism: Rationality, Behavioral Economics, and Public Policy*. Cambridge: Cambridge University Press.
- Sarin, N. (2019), 'Making consumer finance work', *Columbia Law Review*, 119(6): 1519–1596.
- Siipi, H. and P. Koi (2022), 'The ethics of climate nudges: central issues for applying choice architecture interventions to climate policy', *European Journal of Risk Regulation*, 13(2): 218–235.
- Stewart, N. (2009), 'The cost of anchoring on credit-card minimum repayments', *Psychological Science*, 20(1): 39–41.
- Sunstein, C. R. (2017), 'Nudges that fail', *Behavioural Public Policy*, 1(1): 4–25.
- Sunstein, C. R. (2019), 'Ruining popcorn? The welfare effects of information', *Journal of Risk and Uncertainty*, 58: 121–142.
- Sunstein, C. R. (2021), 'Manipulation as Theft', Harvard Public Law Working Papers, No. 21–30. doi:10.2139/ssrn.3880048.
- Sunstein, C. R. (Forthcoming 2023), 'Welfare now', *Duke Law Journal*.
- Tappin, B. M., L. van der Leer and R. T. McKay (2018), 'The heart trumps the head: desirability bias in political belief revision', *Journal of Experimental Psychology*, 146(8): 1143–1149.
- Thaler, R. H. and C. R. Sunstein (2021), *Nudge: The Final Edition*. New York: Penguin Books.
- Thunstrom, L. (2019), 'Welfare effects of nudges: the emotional tax of calorie menu labeling', *Judgment and Decision Making*, 14(1): 11–25.
- Wang, J. and B. Keyes (2014), 'Perverse Nudges: Minimum Payments and Debt Paydown in Consumer Credit Cards', in *Society for Economic Dynamics, 2014 Meeting Papers* 323.