



<http://www.diva-portal.org>

This is the published version of a paper published in *Rationality, Markets and Morals*.

Citation for the original published paper (version of record):

Grill, K. (2014)

Expanding the Nudge: Designing Choice Contexts and Choice Contents.

Rationality, Markets and Morals, 5: 139-162

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-100851>

Kalle Grill

Expanding the Nudge: Designing Choice Contexts and Choice Contents*

Abstract:

To nudge is to design choice contexts in order to improve choice outcomes. Richard Thaler and Cass Sunstein emphatically endorse nudging but reject more restrictive means. In contrast, I argue that the behavioral psychology that motivates nudging also motivates what may be called jolting—i.e. the design of choice content. I defend nudging and jolting by distinguishing them from the sometimes oppressive means with which they can be implemented, by responding to some common arguments against nudging, and by showing how respect for preferences over option sets and their aggregate properties may require the trimming of option sets, as well as helpful choice contexts.

Keywords: Choice Design, Default-setting, Incentives, Nudging, Respect for Preferences.

1. Introduction

Impressed with recent findings in behavioral psychology, both scholars and policy-makers have recently proposed that we should actively design the context in which choices are made, in order to make choices wiser.¹ Such design is usually called *nudging*, after Thaler and Sunstein's 2008 book *Nudge* (new edition 2009). In this article, I endorse nudging, though for somewhat different reasons than those presented by Thaler and Sunstein. I also propose that if we should nudge, we should also *jolt*: If we should design the *choice context* in order to improve the outcomes of choices, we should also design the *choice content* with the same intention. I endorse nudging and jolting as tools, to be used with caution in the pursuit of legitimate aims. I do not endorse them as immune to misuse, universal policy solutions, a third way in politics, or as having any such more

* I wish to thank all fellow participants at the panel on Paternalism and Libertarian Paternalism at the Mancept Workshops in Political Theory in 2013, where an early version was fruitfully discussed. Thanks to Lars Lindblom and Niklas Möller for helpful comments on a later draft. Thanks also to participants of the philosophy seminar at Umeå University, especially Gunnar Björnsson, for constructive comments. Thanks finally to two helpful reviewers for this journal. This research was funded by the Swedish Research Council for Health, Working Life and Welfare, under project 2009-2189.

¹ For an overview over relevant findings in behavioral psychology over the past four decades, see Kahneman 2011. For a number of detailed policy-proposals from academics, see Shafir 2013. On implementations in the USA, see Sunstein 2013.

extravagant properties that proponents have sometimes claimed for nudges, or that critics have taken proponents to claim for them.

Let me illustrate the appeal of jolting with a personal experience: During my graduate years at the Royal Institute of Technology in Stockholm, vending machines were installed in the local subway station. The machines contained bottled water, sweet drinks and candy bars. I remember being a rather frequent customer to these machines from time to time. I was, however, a reluctant customer, experiencing more regret than was usual after similar purchases from manned kiosks. The machines caught me off guard, in a place where I did not expect to be tempted, where I had to stand around waiting for the next subway with nothing better to do than to survey my buying options. There were times when I was genuinely glad to have access to the candy and drinks, but for the most part I wished they were not there.

A nudging approach to subway vending machines would be to arrange the order in which the items are displayed so as to make the least unhealthy products more salient (put them at eye level perhaps). The machines could also be designed so as to be inconspicuous, and information on calorie intake or other health effects could be displayed next to each product. These measures would either activate potential customers' otherwise latent rational deliberation (displaying calories), or circumvent such deliberation (placement). Either way, the content of the choice, that is the available products, would not be affected. The nudging approach would not sanction changing the prices of the products, or, even less, removing any of them.

Nudging is possible because choosers typically enter choice situations without set and stable preferences. The fact that preferences are often indeterminate in this sense can, however, also be exploited by jolts—i.e. by design of the choice content. By adding and removing options, chooser preferences can be influenced in much the same way as by nudging. Removing one option may cause some other option to become more salient, or it may change chooser perspective on the entire choice set. The large table may seem larger if the colossal table is removed. The choice of sitting furniture will become a choice of armchairs if the sofa is removed. Sometimes removing one option entails removing the further options that this one option would open up. With no colossal table, there is no need to choose between enormous tablecloths. Without vending machines in subway stations, there is no need to choose between sweets.

Active design of the choice content, like active design of the choice context, will partly determine what preferences are formed. Therefore, not even the removal of currently preferred options need to conflict with anyone's future preference. For choice design of either type, a choice made under the influence of such design need not be experienced as in any way coerced or hindered. Removing vending machines from subway stations, or not putting them out in the first place, need not frustrate anyone's preference for sweets, even though such preferences are formed once the machines are in place.

2. Nudging—Appeal and Aims

The argument for nudging rests on a rich body of research in the behavioral sciences (most importantly behavioral psychology and experimental economics). It is only because we can make reliable predictions of the behavioral effects of such measures as product placement and information display at point of purchase, on a group level, that nudging is taken seriously as a policy tool. Many findings are disputed and research is ongoing, but some effects are very thoroughly researched. For example, consider the exposure effect: We tend to like things we have been exposed to earlier, regardless of whether we actively remember such exposure, and regardless of whether earlier exposure was combined with a positive experience (Bornstein 1989).

Critics correctly point out that many studies invoked by pro-nudgers concern behavior in isolated cases, and typically in laboratory settings. This is no fault of the studies; it is simply a consequence of the conditions and circumstances of research. Large-scale behavioral change in real life settings is very expensive to produce and difficult to isolate. It may be that the most solid evidence we can find on the likely effects of real life nudging will come from comparative studies of existing differences in choice environments, rather than from controlled experiments (as proposed by Bovens 2012). Obviously, what would make nudging truly appealing is if behavior can be reliably influenced in such a way that there is significant impact on outcomes. It is arguably too early to say to what extent this is possible.

What we can say is that the success of nudging can only be evaluated in relation to the outcomes sought. In formulating their pro-nudging political program, Thaler and Sunstein have incorporated a political aim—to “make choosers better off, as judged by themselves” (2009, 5). It seems this aim is proposed in large part in order to avoid controversy (which would be in line with Sunstein’s long-time preference for “incompletely theorized agreements”—see Sunstein 1995). However, the justification that “we are just giving people what they want” has never been unproblematic in political contexts. It is all the more problematic in light of the behavioral research that underpins nudging. One of the best known results in behavioral psychology is that preferences vary with how alternatives are described, even when descriptions are short and clear and have logically equivalent content (McNeil et al. 1982). Moreover, even if it were unambiguous what every person judged to be good for her, in every regard, it would be controversial how to aggregate these judgments or preferences into a social preference, which we would have to do for any nudge that affects groups rather than single individuals. Such aggregation issues are explored in social choice theory. Furthermore, even if individual judgments were aggregated in some uncontroversial manner, it is far from uncontroversial that we as a society should strive to promote our own wellbeing as judged by ourselves. Perhaps we should rather strive for excellence, or perhaps we should make it a priority to preserve the natural environment, or our own cultural artifacts.

Thaler and Sunstein say very little on the aim of nudging. More generally, pro-nudgers tend to sidestep the issue of the precise goal of nudging and so whether this goal can be effectively achieved. This may be because Thaler and Sunstein have presented two arguments that, if valid, indicate that it does not matter much what nudging gives us, as long as it gives us something: The first argument is that nudging is unavoidable. The second argument is that nudges can easily be resisted and so pose no threat to liberty. If either of these arguments is sound, whether we should use nudges on any one occasion would seem to come down to straight-forward cost-benefit analysis, even if the input to such analysis may be uncertain.

I propose that the argument from unavoidability should be accepted after modification. I propose that the argument from preserved liberty should be rejected, for several reasons. I will develop and defend these claims in the following two sections. The upshot is that in light of our increasing knowledge of the behavioral effects of choice contexts, we have good reasons to design contexts so as to promote legitimate aims. This leaves it an open question what these legitimate aims are. I will not take a stand on this important question in this contribution, but will rather assume that governments and other actors whose decisions structure our choices can have legitimate aims and that, when they do, these aims can be furthered by using nudges, as well as jolts.

3. Unavoidability and Moral Neutrality

The argument from unavoidability is arguably Thaler and Sunstein's most important argument for nudging. It is the argument they focus most on in their 2003 article (Thaler and Sunstein 2003, 1171–1183). They introduce the argument early on in *Nudge*:

“In many situations, some organization or agent must make a choice that will affect the behavior of some other people. There is, in those situations, no way of avoiding nudging in some direction.” (Thaler and Sunstein 2009, 5)

Thaler and Sunstein return to this argument frequently in the book, e.g. noting in a late chapter that it has been “emphasized throughout” (2003, 240–241). In his recent follow-up book *Why Nudge?*, Sunstein again emphasizes this argument, saying for example that “it cannot be repeated often enough [. . .]: *Choice architecture is inevitable.*” (2014, 118, original emphasis)

The argument from unavoidability is based on the straightforward observation that whenever something is designed, it will inevitably nudge people towards some behavior or other, whether or not this is intended by the designer or designers. What is designed may be a building, a search algorithm, or health care reform. Whatever it is, the design will facilitate some behaviors and make alternative options less salient, less attractive, or slightly less accessible. The

indication from behavioral science is that these features of the design will have substantial effects on behavior in the aggregate. The argument from unavoidability is that since nudging is both unavoidable and efficacious, we might as well ensure that it is beneficial.

One weakness of this argument is that it fails to distinguish between nudging that is intended to shift behavior in some particular direction, and nudging that has no such aim and is in that sense unintended. Thaler and Sunstein explicitly note that nudging can be either intended or unintended (or “unintentional”). They downplay the distinction, since, in either case, “nudges can have major effects” (Thaler and Sunstein 2009, 10). However, what critics are concerned with is typically not that people are influenced by their choice contexts, but that their choice contexts are designed, by others, with the intention of producing some particular behavior.²

Intended nudging is not unavoidable. As Thaler and Sunstein note (2009, 10), it can be avoided by randomizing between design alternatives, though this seems silly. Arguably, it can also be avoided by letting design decisions be made on the basis of other things than behavioral effects, such as cost minimization or aesthetic concerns. Consider the decoration of some large public area, such as a square or a park. This project may offer opportunities to nudge people towards exercise by placing exercise machines in strategic locations, or to nudge people towards healthier lifestyles by displaying persuasive messages or pictures. However, such nudging is easily avoided. Design decisions can be determined by costs and aesthetics only—the most beauty for the buck. Different designs may still affect behavior differently: A Zen garden may inspire meditation and large open areas may inspire sport and play. However, if these effects, though perhaps foreseen, do not provide motivating reason either for or against any particular design, then I believe nudge skeptics will be silent.³

While intended nudging is not strictly speaking unavoidable, the argument from unavoidability is strengthened by the fact that our knowledge of behavioral psychology is steadily increasing and spreading. Once choice designers have this knowledge, or even just the knowledge that such knowledge exists, their decision problems change irrevocably: Enlightened choice designers must choose between either using findings from behavioral research to influence behavior in various directions, or actively avoiding such influence. Avoiding will become more and more difficult as behavioral insights are disseminated through the population. Large institutional choice designers like governments can com-

² For example: “[Nudges] are manipulative, first because the government employs them *with the intention* of affecting people’s choices.” (Grüne-Yanoff 2012, 636, my emphasis)

³ The role and moral significance of intentions is of course a much debated issue in moral philosophy and there are several fine-grained distinctions. My analysis only presupposes the distinction between on the one hand straight-forwardly intending something and, thereby, causing it, and on the other hand causing something as a mere side-effect. This distinction is familiar from the discussion on the controversial doctrine of double effect (for an overview, see McIntyre 2011, for a critical discussion, see Scanlon 2008, e.g. 22–23). Importantly, I make this distinction only to identify controversial forms of nudging. I do not thereby endorse the moral relevance of the distinction.

mission design according to formal requirements that include that behavioral effects not be considered in design decisions. However, such requirements will be difficult to monitor. Enlightened individual designers may still allow behavioral effects to be decisive, especially when other criteria are not at issue or very marginally affected. For example, the park designer may put benches in groups, rather close and slightly facing each other, in order to promote socialization, thinking that people in the area are alienated or lonely and would benefit from more interaction. This arrangement may also be aesthetically pleasing and no more costly than alternative arrangements.

Despite these considerations to strengthen the argument from unavoidability, I believe that we must concede to the critics that *intended* nudging can often be avoided, and sometimes quite easily. I propose, however, that whether or not nudges are intended is not morally relevant. One way to phrase this is to say that the mere intention to influence behavior via choice context is not morally problematic. Another way of phrasing it is to say that choice contexts are not morally neutral just because they have come about unintentionally. I therefore call this argument for nudging the *argument from non-neutrality*.

It may be that some behaviors are morally problematic, or impermissible, if and only if they are intended. Clearly, however, merely influencing other people's behavior via design of their environment is not one of those behaviors. Consider putting up a warning sign by a power station, with the effective intention to cause people to avoid trespassing, or be more mindful of where they are going. This is not morally problematic. Therefore, if intended nudging is morally problematic in a way that unintended nudging is not, there must be something about nudging beyond its influence on behavior that explains this difference. I will consider and tentatively reject some such possibilities below.

At this point, I will bluntly declare that I do not believe that even intended nudging, that is influence on choice and behavior *via choice context*, is by itself morally problematic. If I pick one of two available warning signs because I know its design is more likely to induce caution, this is in no way morally problematic. If the park designer opts for yellow garbage cans in order to reduce littering, this is equally unproblematic.

There are of course, for any agent, many things it may not do. Perhaps not anybody has the right to put up that warning sign. Certainly not everyone has the right to paint public garbage cans yellow. Perhaps the government has no right to interfere with certain private exchanges. However, when some agent has a general jurisdiction or right to shape or regulate some environment, there is no general moral problem with doing so with the intention to influence choice and behavior. What may be problematic is for an agent to influence behavior with an aim that is not among that agent's legitimate concerns.

I do not want to deny that there may be a special quality to the complete absence of design. Perhaps one quality of untouched wilderness is that it does not embody other people's attempts to influence us in any way. If that is so, the same could perhaps be said for artificial environments that are created without any consideration for how they may influence us. Consider for example environ-

ments we encounter when we travel to places that are built for people in other circumstances and that have not (yet) adjusted themselves to visits from people like us (i.e. to tourism). While I grant that these are possibilities, I propose that if an environment is artificially designed for us, as most environments are for most people living in them, then avoiding choice context design does not add any special value to that environment.⁴

I conclude that, at least in artificial environments, there are no morally neutral choice contexts. This implies that, for any choice context, if it can be costlessly changed, leaving it unchanged requires justification in just the same way as does changing it. A new choice context may bring greater risks than an existing one or it may upset legitimate expectations, but this cannot be taken for granted in any one case. Sometimes avoiding risks and meeting legitimate expectations favor changing a choice context. Consider for example a change in the default employment contract to include a retirement savings plan.⁵ This change seems to reduce long-term risks and to honor employees' legitimate expectations of their employers or of their government to care for their long-term financial wellbeing.

The argument from non-neutrality is more robust than the argument from unavoidability because intended nudging is not, strictly speaking, unavoidable. The argument from non-neutrality is also stronger because that nudging is unavoidable is consistent with its being morally neutral, and so requiring no justification. This would imply, for example, that in the face of controversy over changes, choice contexts could unproblematically be left as they currently are. Because no choice contexts are morally neutral, however, we should prefer that context that is most conducive to the legitimate aims of the choice designer, unless there is some other moral issue at stake. Again, I will consider some such moral issues below, after I introduce choice content design. Before I even do that, however, I will consider and reject Thaler and Sunstein's second main argument for nudging.

4. Significant Incentive Effects

Thaler and Sunstein's second main and recurring argument for nudging is that since it does not significantly affect incentives, it poses no threat to liberty. Nudging is, as they put it, "freedom-preserving", and this in contrast to many other measures the government can take to promote welfare.⁶ There are two

⁴ This is not to deny that there are down-sides to the comprehensive reliance on design throughout society. Such reliance might cause the loss of spontaneity and of chance discoveries, including discoveries of new legitimate aims. On this, see Dworkin 1988.

⁵ On the effects of such rules on enrollment, see Beshears et al. 2009.

⁶ Thaler and Sunstein say "liberty-preserving" in *Nudge* (2009, 5) and Sunstein says "freedom-preserving" in *Simpler* (2013, 2). Like them, I make no distinction between these concepts in this context. It is because nudging is allegedly freedom-preserving that the pro-nudging position has been called 'libertarian paternalism'. This label is contested, for good reasons. It completely

steps in this argument: 1) Nudges do not significantly affect incentives. 2) Measures that do not significantly affect incentives preserve freedom. I will argue that both steps are unwarranted, for independent reasons.

Before evaluating the freedom-preservation argument, I must note a conceptual complication relating to step 1: Many would consider it a necessary condition for some measure to be a nudge that it does not significantly affect incentives. Indeed, this condition is generally taken for granted in the debate. For example, noted critics of nudging Hausman and Welch sum up their understanding of the concept so: “Nudges are ways of influencing choice *without limiting the choice set* or making alternatives appreciably more costly in terms of time, trouble, social sanctions, and so forth.” (2010, 126, my emphasis) If this is right, it is incoherent to question whether nudges significantly affect incentives or not. Moreover, some would even consider it a necessary condition for some measure to be a nudge that it is freedom-preserving. For example, Yashar Saghai’s development of the concept explicitly incorporates this condition (2013, 487). Having noted these conceptual complications, I will persist in treating the claim that nudges are freedom-preserving as an argument for the moral innocence of nudges, rather than as a semantic thesis about what defines a nudge.

Step 2 of the argument can be evaluated independently of the concept of nudging. I propose that this claim does not hold up to scrutiny. Whether or not some measure preserves freedom depends, of course, on what freedom is. Without attempting to summarize the contemporary debate on the nature of freedom, I will simply claim that here is strong consensus in this debate that the imposition of costs limits freedom in one way or other. If an option is rendered more costly, I am less free to choose it. At least if this option is my preferred option, I am thereby rendered less free. The option is, as Philip Pettit puts it, replaced by a “burdened counterpart” (2009, 11). For another example, consider Isaiah Berlin’s view that a person’s freedom is “a function of what doors, and how many, are open to him; upon what prospects they open; and how open they are.” (2002[1969], 41) Doors are more open if they are less costly to take. There are, of course, exceptions to the consensus. Thomas Hobbes seems to have held the view that you are free to do something as long as you are physically able to do so, regardless of the costs involved.⁷

Assuming that costs typically do limit freedom, is there some threshold of significance under which they fail to do so? I know of no principled argument to this effect. It may for practical reasons be pointless or pedantic to insist on the loss of freedom when it is very small, but I see no theoretical reason to

disregards the rich libertarian tradition built around strong ownership rights. ‘Libertarianism’ in the present context only means the priority of freedom of choice over other values. I will not engage further with this terminological issue.

⁷ Some contemporary authors hold that one is free to do whatever one is physically able to do, regardless of costs, but they then also hold that imposed costs diminish one’s overall freedom (Steiner 1994; Carter and Kramer 2008). Pettit is an exception in that he holds that costs only diminish freedom if they are imposed arbitrarily, but then he holds that this goes for any measure, including outright compulsion (e.g. Pettit 1997).

disregard it.⁸ We may think that particular freedoms, such as the freedom of speech, are preserved even if their exercise is costly, in the sense that people are sufficiently free to speak (sufficiently to claim that democratic rights are respected etc.). However, it is unwarranted to claim quite generally that freedom is preserved if costs or burdens are small. At the very least, this claim requires some specification as to what defines the threshold and why smaller costs are irrelevant from the perspective of freedom. Thaler and Sunstein offer no such specification or argument.

If light of these observations, I propose that even if nudges were “easy and cheap to avoid” (Thaler and Sunstein 2009, 19), this would not by itself imply that they preserve freedom. Many nudges limit freedom only very little, but it is important not to confuse small costs with no costs. This is particularly important since nudges are intended to be used extensively and since a series of small costs may be large in the aggregate. Changing a default rule from opt-in to opt-out may impose only very small costs on choosers who will actively opt out under the new rule. However, with a growing number of such changes, costs may not stay low. Choosers who are consistent outliers may have to constantly opt out of arrangements set up to benefit the typical chooser.⁹ Sunstein has noted this potential problem and proposed in response that default rules may be personalized (2014, 99). However, this does nothing to strengthen the claim that small costs do not limit liberty.¹⁰

In evaluating step 2 of the freedom-preservation argument, we must keep distinct the independent argument from non-neutrality (and the argument from unavoidability). Some features of choice contexts are unavoidable. For example, there must often be a default rule, and there is typically a cost to opting out, whichever option is the default. The opt-out cost is not neutral under any particular default, and so it is both physically and morally unavoidable that there be some opt-out cost. This may imply that the setting of a default does not diminish freedom.¹¹ If it does imply this, however, this may be because the cost is

⁸ Libertarians in particular are typically unwilling to accept even limited state interference with voluntary free market transactions that do not harm third parties. Distortions can be larger or smaller, but they are distortions all the same. This is another reason for why the term ‘libertarian paternalism’ is so contentious.

⁹ Sarah Conly (2013) considers the burdens on consistent outliers a valid objection to coercive paternalism, though not decisive, since these burdens may be outweighed by benefits to non-outliers (64–66).

¹⁰ None of this is meant to deny that some nudges may impose no costs at all. Using different sorts of framing effects may affect the salience of different options without making any alternative more costly.

¹¹ More specifically, freedom is not diminished by setting a default if freedom is about what options are available to us at what cost. It may diminish freedom if freedom is about what causes us to have the options we have. For example, some hold that freedom can only be diminished by intentional action. On such a theory of freedom, changing a default rule may diminish freedom even if the new default rule is welcomed by choosers, provides a richer set of options, and improves chooser wellbeing. For what it is worth, I find this implication an embarrassment for such views of freedom.

unavoidable, not because it is small. If so, it is the argument from non-neutrality that is doing the work.

For an example, assume a work-out machine is available for employees at some workplace. Consider a rule that requires employees to fill out a brief lifestyle questionnaire before using the machine. The intention may be to nudge towards health-promoting lifestyle changes (unrelated to safe use of the machine). This requirement entails an added cost of using the machine. Assuming this cost is insignificant, it is the right sort of cost to consider in evaluating the freedom-preservation argument. However, consider now that the machine has two different settings, for doing two different exercises, and that changing between settings is a bit cumbersome. When not in use, the machine automatically returns to the default setting. The employer picks a default. For users who prefer the non-default setting, switching is an added cost of using the machine, perhaps just as large as filling out the lifestyle questionnaire. However, given that preferences vary in the group, and given how cumbersome it is to change the settings on this machine, it is unavoidable that some users face this cost. If this sort of cost does not diminish freedom, this is likely due to unavoidability, or non-neutrality, rather than to the smallness of the cost.

If I am right in rejecting step 2 of the freedom-preservation argument, step 1 on its own will not make an argument. However, I may not be right. Also, step 1 is of interest either way, for two reasons. First, and most obviously, if nudges can have more than insignificant incentive effects, this will make it (even) harder to argue that they are freedom-preserving. Second, less obviously, the claim that nudges have only insignificant incentive effects entails further problems for the pro-nudging position.

Consider the setting of a default rule that will affect 100 people and where there are two possibilities:

Risky default: Everyone can opt out at insignificant cost. 89 persons will do so. 11 persons will not and so risk serious harm.

Safe default: Everyone can opt out but the cost is significant. One person will do so. This person will risk serious harm.

If step 1 is correct, then setting the risky default will be a nudge, while setting the safe default will not be, since with this default the opt-out cost is not insignificant. This may be problematic for two independent reasons: First, default-setting is a paradigmatic example of a nudge, as well as of the unavoidability of choice context. If some default-settings are not nudges, this would not only raise questions regarding the concept of a nudge, it would seriously weaken the argument from unavoidability: If people are not nudged by the safe default, then we can easily avoid nudging in this case, contrary to Thaler and Sunstein's unavoidability claim. Note that this point is independent of intent: The safe default is no nudge, whether or not it is set intentionally.

Second, the safe default seems very attractive in the example. The opt-out cost for the one person may be significant for her in some sense—i.e. it may cost her some real money, or substantial effort. This cost, however, seems rather

unimportant in relation to the risk of serious harm to ten people. Especially so since there need not be any benefit to the ten, and they may have no independent or stable preference for the risky option. Their failure to opt out may simply be due to inertia, misunderstanding, or any other non-rational influence. This is not to mention that the individually insignificant cost to the 89 people who opt out under the risky default may in the aggregate be comparable to or greater than the one significant opt-out cost. All this means that if nudges do not significantly affect incentives, there seems to be little reason to prefer nudges over non-nudging default-settings. Even more to the point, it seems highly questionable whether nudges are more freedom-preserving than non-nudging default-settings.

To further illustrate my point, consider this possible story behind the two default rules: In a public park on a mountain, there are two foot paths. One is rather dangerous but has a spectacular view. The other one is safe but has a slightly less spectacular view. A ranger hut (or something) must be placed at the entrance to the park, where both paths begin. Because of the difficult terrain (and the low budget) there are only very limited options for placing the hut. It can either be placed so that it will hide the safe path from view, or it can be placed so that it blocks the dangerous path entirely, making it accessible only via a long detour. Assume that information on the two paths is readily available, perhaps from a large map on site, and that the dangerous path is marked with a clear warning sign. Still, with the safe path hidden from view some people will take the dangerous route. Most of them will do so not because they consider the view worth the risk, but because they are absent-minded and do not pay much attention to maps and signs. They are intent on hiking and they will just not bother to look for the safe path. These are the ten people who will (happily) take the safe path under the safe default, but will (inadvertently) take the dangerous path under the risky default. One person will take the dangerous path even if she must take the long detour, perhaps because experiencing spectacular views is just that important to her.¹²

At times, Thaler and Sunstein indicate that their position is not that nudging has or should have only insignificant incentive effects, but rather that opt-out costs should be kept low whenever possible. For example: “Our goal is to allow people to go their own way at the lowest possible cost.” (2009, 252) This position would be consistent with preferring the safe default in the example and in general to prefer the best default all things considered, irrespective of the size of the (minimal) incentive effect. However, to take up this position would also be to give up the freedom-preservation argument.

In conclusion, both steps in the freedom-preservation argument are deeply problematic. As noted in the first paragraph of this section, the limited impact on incentives is generally considered part of the essence of nudging. If we reject the freedom-preservation argument, and in particular its first step, where does

¹² Another possibility is that not ten but eleven persons are absent-minded and eager hikers, while one person would be triggered by the blocking of the dangerous path and take it just because it is blocked (“no one is going to stop me from hiking wherever I want”).

this leave the concept of nudging? I propose to give up on limited incentive effects and define nudging only in terms of effects on choice outcome from choice context, or, as Thaler and Sunstein prefer to say, from choice architecture. We can then say that some nudges are freedom-preserving and that this is a virtue of those nudges. However, as pro-nudgers we should not hold that all nudges have this virtue. If my proposal is conceptually too radical, the alternative is to keep the insignificant incentive effects condition on nudging, and instead give up some common claims on the virtues of nudging relative to other influences on choice.

5. Choice Content Design

I have argued that the non-neutrality of choice contexts indicates that nudging, i.e. choice context design, can and often should be used to promote legitimate aims. I have postponed discussion of some counter-arguments in order to simultaneously consider those arguments as they apply to both nudging and jolting—i.e. to choice context design as well as to choice content design. In this section, I will present my *prima facie* case for choice content design. I will explain how, just like nudging, jolting has non-rational influence on choosers. In the subsequent section, I will argue that relevant counter-arguments to jolting also apply to nudging, and so that if we should nudge, we should also jolt.

One criticism of nudging is that it is no new policy tool, but merely a new name for established methods of changing and steering behavior by adjusting the living environment so as to “make healthy choices the easy choices”, as the public health pioneer Nancy Milio put it already in 1976 (Quoted from Vallgård 2011, 201, who develops this critique). This critique is warranted in light of some exaggerated claims to novelty, but I believe the concept of nudging, and the discussion it has spurred, has helped many of us realize that all choices have contexts, and that these contexts can be designed so as to promote some outcomes at the expense of others. To use nudging as a policy tool partly means, I propose, to adapt a perspective on choice contexts as ever-present and designable. It means, furthermore, to consider, in light of behavioral science and the peculiarities of the choice at hand, what designs could promote legitimate aims. It means, finally, to implement these designs. Established policy tools such as social advertising certainly aim to change choice contexts. However, I propose that it is a relatively new or at least under-utilized idea to perceive the context of a choice in its entirety as an object of design.

Similarly, to use choice content design as a tool means to adapt the perspective that choice contents are ever-present and designable. It is to perceive the creation, regulation and shaping of choice contents as a design problem. It is, specifically, to pay attention to the many ways in which choice content affects what is chosen. Such effects are not limited to the obvious circumstance that options that are not accessible will not be chosen. Behavioral science has shown

that just as choice contexts can affect what is chosen in surprising and arguably non-rational ways, so can choice contents. The most well-know example of this is probably shifting preferences among options as they are complemented with additional options. Seemingly rational people who prefer A to B in the set {A, B} may prefer B to A in the set {A, B, C} (see e.g. Tversky 1972). The rationality of such shifts in preference is much discussed in normative decision theory. Their reality, however, is undisputed.

I will now explain and defend choice content design by considering some variations on the now famous (in some quarters infamous) cafeteria case, introduced by Sunstein and Thaler (2003, 1164). Like the original case, mine is hypothetical. In expanding on the case, I will invoke psychological explanations of choice behavior that I hope will be both intuitively plausible and consistent with main trends in current psychological research.

Let us, like Sunstein and Thaler, “consider the cafeteria at some organization” (ibid.). In order to avoid special concerns with children or minors, let the cafeteria be one for adult employees.¹³ For some reason, this cafeteria must serve deserts (perhaps employees demand it). However, the cafeteria director has considerable leeway to decide what desserts to offer and how to present them to customers.

Assume that the status quo is serving six desserts, A–F. That this is the status quo may mean that these are the desserts that are offered at present, or it may mean that these are the desserts that are available from the current supplier, or it may mean that they are the desserts the director first comes to think of. In order to discuss the case, I need to assume a ranking of these options in terms of some legitimate aim. Assume, therefore, that promoting employee health is a legitimate aim for this cafeteria director, perhaps by mandate from central management. In terms of this aim, the desert options are ranked from A, which is best, to F, which is worst.

Let us now consider, true to liberal fundamentals, an individual dessert-craving costumer. In the status quo, this customer will opt for F. However, by changing the order in which desserts are displayed, the director can cause the customer to opt for E. This may be achieved, for example, by placing E earlier in the line or giving it a more pronounced placement. This is nudging.

The director cannot, by any sort of choice context design, cause the customer to opt for any of A–D. However, the director can also design the *choice content*: By introducing a new option G, the director can cause the customer to opt for D. G may be, for example, an option that is even less healthy than F, and more obviously unhealthy, making the costumer conscious of the health effects of desserts and so causing her to opt for something less unhealthy. Moreover, by removing the options A and B, the director can cause the costumer to opt for C. This may be, for example, because the customer is attracted to the least unhealthy option given that it is not too healthy (and so not too dull).

¹³ Thaler and Sunstein introduce this complication in *Nudge*, seemingly unnecessarily, by changing their example relative to their earlier rendering.

These two instances of choice content design are different in that one widens the scope of choice and the other restricts it. We may say that the first increases freedom, while the second limits freedom. To limit freedom is often seen as illiberal, and correctly so, but this judgment should not be made too hastily in any one case. More choice is not always better than less, nor is it always preferred by choosers. More options generally make choosing more cumbersome and make it more difficult to identify one's favored option. More options also increase the risk of making a mistake in choosing, and if the extra options are poor they raise the stakes involved in such mistake. All this may make more choice both less desirable and, sometimes, less desired.¹⁴

Furthermore, we should not be unduly influenced by the fact that I have presented the design problem in relation to the choice set A–F, assumed to be the status quo. This status quo may well be morally arbitrary. I offered as a possible explanation for this status quo that A–F might be what the current supplier has to offer. However, there may be no reason to favor this supplier (there may even be reason to favor a different supplier). More generally, it is often the case that some agent, government or otherwise, must design a choice from scratch, with no status quo to defer to. When the cafeteria opens for the first time, there may be only two sorts of constraints: The requirement that desserts should be available, and practical constraints such as limited fridge room and the financial incentive to avoid waste. Now, should the director offer five or ten or twenty different desserts? A higher number is clearly not always better, even for customers—just consider the inconvenience of having to choose between five hundred different desserts, displayed along a long line, or listed on a long menu. An adaptive choice mechanism such as a search tool with dynamic categories could mitigate the downsides of too many options, but that just goes to show that the number of options should not be maximized, but instead adjusted in relation to other features of the choice environment, such as searchability, as well as to what values are at stake.¹⁵

Returning to our hypothetical cafeteria director, the choice content designs already discussed—adding an unhealthy G and removing the relatively healthy A and B—do not exhaust his options. The director can also remove F, causing the customer to opt for B. This may be, for example, because F is quite tempting and so with F out of sight (and smell), the customer's willpower is more intact and so she opts for a rather healthy dessert.

¹⁴ On these and other downsides of choice, see Dworkin 1982. It has been observed that too many options cause customers to refrain from choosing. For a recent economic model for preferences over choice sets with varying numbers of alternatives, see Kuksov and Villas-Boas 2009. For readers concerned that choices be rational, note that removing options can cause one's set of preferences to be more coherent. As noted above, even seemingly rational people may prefer A in the set {A, B} and yet prefer B in the set {A, B, C}. By removing C from the latter set, this incoherence (if it is one) is also removed. In the status quo, of course, people may not even be seemingly rational.

¹⁵ The status quo need not be morally arbitrary of course. If the cafeteria has been offering a certain set of options for some time, and customers have grown accustomed to having those options available, it may frustrate legitimate expectations to remove these options.

F is the option the customer would have chosen in the status quo. It may therefore seem particularly problematic to remove this option. However, I have concluded that the status quo is often morally irrelevant. ‘Removing’ is really a misnomer. We should rather talk of non-inclusion. It may perhaps still seem particularly problematic not to include options that would have been chosen if they were included. However, F is not the only such option. B is chosen from the set {A, B, C, D, E}. Relative to this set, non-inclusion of A and B and inclusion of F will cause a choice of E or F, depending on choice context. Does this mean that non-inclusion of B, E or F is particularly problematic? It does not. Almost all options would be chosen from some choice set. Even the tasteless A would be chosen over the directly harmful Z, or the even more tasteless A'. This does not, of course, imply that all options should be included in the actual choice set.

My cafeteria example shows how choice content can influence choice in other ways than by blocking a choice of some option by removing it, or causing a choice of some option by adding it. Some or all of the hypothetical influences may appear non-rational. Consider how the mere presence of G makes the value of health more salient to the customer. On the face of it, it does not seem rational to value health higher because an additional option is present. On the other hand, guarding against unhealthy choices is more important the more unhealthy are the available options. It is often hard to say what is rational and what is not. I take no stand on such issues. The point is that there are diverse and intricate ways in which choice content determines choice. Depending on the circumstances, these effects can to some extent be predicted and so choice content designed on the basis of such predictions.

6. Arguments Presented against Nudging

As stated in the very first paragraph of this article, I endorse nudging and jolting only as tools to be used with caution and in the pursuit of legitimate aims. Importantly, I do not endorse libertarian paternalism. I emphasize this, since Thaler and Sunstein have coined both expressions and consistently present them as intimately interconnected. However, and Thaler and Sunstein do not deny this, libertarian paternalism is the idea that nudging should be used for particular purposes (the promotion of subjective wellbeing discussed above) and for particular reasons (including the freedom-preservation argument that I have rejected). Naturally, these purposes and reasons have attracted criticism that is not relevant to nudging as such.

Nudging should be distinguished not only from its various possible justifications (and their controversial names), but also from different means that may be used for its implementation. Like any tool, nudging can be misused. Thoughtful descriptions of alternatives is typically a non-intrusive nudge: One of two equally concise, clear and relevant descriptions can sometimes yield better health outcomes than the other. However, alternatives can also be described

with reference to completely irrelevant properties, in a manner that simply confuses choosers, or choosers can be overwhelmed with too much relevant information. Such misuse must not be confused with the tool itself.

In the very same way, we should distinguish between inherent problems with choice content designs and contingent problems that may or may not be associated with such designs. Such problems may be more or less directly moral. They may concern duties or rights or they may concern consequences in a wider sense. Non-inclusion of an option that the chooser has a right to will of course imply a rights violation. Removing an option by criminal prohibition is generally very costly and very intrusive (see e.g. Husak 2013). For both of these sorts of reasons, some authors oppose criminalization of unhealthy or otherwise harmful behavior, but favor non-legalization in the form of prohibitions on the manufacturing and distribution of the means to such harmful behavior, or otherwise the facilitation of this behavior (e.g. De Marneffe 2009, concerning prostitution). As John Stuart Mill proposed, though the government has no right to decide, “for purposes of repression or punishment”, that some conduct should be prohibited, it may be right “in endeavoring to exclude the influence of solicitations which are not disinterested, of instigators who [...] have a direct personal interest on one side” (1859, chapter V, 8th paragraph). At least Mill found “considerable force in these arguments” (ibid.). These are complicated matters. The point is only that a choice set can sometimes be limited quite easily and at little or no cost, as when the cafeteria director decides against adding another dessert to the menu, or for subtracting one previously on it, while other times such limitation of options is much more difficult and can only be achieved via measures that are in themselves oppressive or illiberal. To endorse choice content design is not to endorse such oppressive measures.

With these distinctions in mind, I will turn now to some arguments that target nudging understood as a tool.¹⁶ In the previous section, I tried to rebut two arguments against jolting specifically: That it limits freedom and that it removes options that would otherwise have been chosen. I expect that an opponent of nudging would be likely to consider the arguments I will now discuss to tell against jolting as well. The recent literature on arguments against nudging, and their refutations, is very large. I can only very briefly consider what I take to be the three most frequently voiced arguments: That (intentional) nudging is deceitful or non-transparent, that it is manipulative, and that it prevents moral development. Though it would be useful to consider these arguments in further detail, I will limit myself to pointing out some weaknesses in them and to explain why they are no stronger against jolting than against nudging. This will serve to strengthen both of my theses: That we should nudge, and that if we should nudge, we should also jolt.

¹⁶I will not consider general arguments against involvement in other people's lives, especially by the government. There are of course legitimate concerns with the incompetence and corruption of people in the government and elsewhere, but these concerns are too general to warrant treatment here.

The claim that nudging is deceitful has an empirical part and a normative part. The empirical part is that nudges often or always work best, or only work at all, when they are covert—i.e. when nudgees do not know they are being nudged. Claims to this effect are typically unsubstantiated.¹⁷ One recent study indicates that default-setting has substantial effects even when people are informed immediately prior to their choice that they have been defaulted into one of the options.¹⁸ More generally, it seems psychologically very likely that whether or not being informed that one is being nudged has any effect on the impact of the nudge depends on whether nudgees have any resistance to being nudged.

The normative part of the argument is that it is deceitful not to inform nudgees that they are being nudged. Luc Bovens (2009) has developed a version of this claim in some detail, formulating it in terms of transparency. He proposes that nudges, in order to be morally acceptable, should be possible to discern by all nudgees (216–217). It is not clear what is the moral basis for this requirement. Bovens mocks political transparency, assuming such transparency can only be general and not specific. However, political transparency typically entails that information on policy is made available via established channels, easy to find for those who seek it. Such information can be very detailed. Extensive and detailed transparency (including what Bovens calls “token transparency”) is consistent with lack of local discernibility. I propose that local discernibility is far too demanding a requirement, especially considering that our living environment is the result of countless actions by other people, organizations and institutions, many of which aim to influence our behavior one way or other.

Failing to see the force of the argument from deceit or lack of transparency, I speculate that it is fueled by some underlying worry. Perhaps by this two-fold concern: First, we may be nudged, and jolted, according to someone else’s agenda. This may be problematic in itself, and especially so if the agenda is not in our best interest, or not what we prefer, or not what we would prefer after due informed consideration. Second, nudges, and jolts, can be plentiful, subtle, and individually relatively insignificant.¹⁹ This means that we will not pay attention to them, will not be motivated to try to counter them, and may not even have sufficient reason to do so in any one case. The result can be that we are under the potentially malign influence of other people, with no obvious way

¹⁷ For example, Adam Burgess claims that “nudges are likely to work best precisely when they are not transparent” (2012, 11). Burgess provides no support for this claim, but only refers to Luc Bovens 2009, who first speculates that if we let people know they are being nudged, “the intervention may be less successful” (209), then simply asserts that “these techniques do work best in the dark” (217).

¹⁸ In a presently unpublished manuscript accessible through the Social Science Research Network, George Loewenstein, Cindy Bryce, David Hagmann and Sachin Rajpal conclude: “Pre-informing people of defaults did not have a measurable impact on their effectiveness.” (Accessed 28 march 2014 at <http://papers.ssrn.com/abstract=2417383>)

¹⁹ Somewhat similarly, Sunstein speculates in his most recent book that the worry with transparency is really a worry about the limited salience of nudges. Sunstein rejects this worry as unfounded since many nudges are quite salient (2014, 148–149).

out. More direct and more harmful influences are typically more easily noted and countered. If I am right in my speculation, the transparency and deceit worry is actually a worry about manipulation.

In the nudging debate, explicit arguments from manipulation do not have this form. Instead, they point out that nudges exploit our mental weaknesses (Bovens 2009, 209; Goodwin 2012, 86), or that they “deliberately circumvent people’s rational reasoning and deliberating faculties” (Grüne-Yanoff 2012, 636), or, similarly, that they are “circumventing the individual’s will” (Hausman and Welch 2010, 130). If this is manipulative, jolting is manipulative in the same way. However, I fail to see why it would be necessarily problematic to manipulate people in the sense of circumventing their rational will. As Sarah Buss has argued, we are constantly subject to nonrational influence: “No rational chooser can do anything without the aid of nonrational influences that determine how she sees her choice situation and how she weighs the options she sees.” (2005, 214) That we are subject to such influence does not mean, Buss convincingly argues, that we cannot act for reasons, and for our own reasons. Moreover, whether any particular circumvention of one’s rational self is or is not problematic would seem to depend in part upon whether or not one welcomes such circumvention (as has been proposed, for the case of nudges, by Wilkinson 2012, 13).

Perhaps, like the worry with deception, the worry here too is that, when our rational selves are circumvented, we can become the instruments of someone else’s will.²⁰ The two worries can reinforce each other—influence that circumvents our rational selves may be harder to notice and counter. The issue of when one agent becomes the instrument of another agent’s will is extremely rich and complex, being central to such moral issues as the distinction between coercion and manipulation, the nature of authenticity and of moral responsibility, and the existence or not of a free will. I will have to leave this issue to one side, noting only that if this is an objection to jolting, it is to the same extent an objection to nudging.

The third and final argument that I will consider is that, as Adam Burgess puts it, “nudging crucially limits the possibility of conscious learning and improvement” (2012, 16). Bovens similarly worries that “the long-term effect of Nudge may be infantilisation” (2009, 214–215). On the face of it, there seems to be no particular connection between the extent to which the choice context is designed to promote legitimate aims and the extent to which people develop morally. Sarah Conly considers this issue in detail concerning coercive paternalism, noting that “rather than weakening people’s ability to make good decisions, paternalistic legislation can help habituate them to making good decisions”.²¹

A related worry that may be underpinning the worry over infantilization is that behavior supported by nudges will not be resilient. Thus, Bovens claims, nudging “may leave the agent with a lack of moral strength to implement the target behavior once the regulation or the environment cues are no longer present”

²⁰ Cf. Michael Garnett’s recent (2013) defense of self-rule as absence of rule by others.

²¹ In a presently unpublished manuscript, Andres Moles develops an analogous argument for the case of nudging in particular.

(2009, 24). Conly, drawing a parallel to public education, counters that we are dependent on a great many things, and that if these things promote good outcomes, our dependence on them is an argument for making them as permanent as possible, rather than for avoiding them (2013, 68–69).

As Bovens admits, it is largely an empirical question what are the likely effects of nudging when it comes to moral development and the resilience of positive effects. I suspect that many concerns are due to the relative novelty of nudging and the lack of familiarity with the tool among the general public. While prohibitions are not novel in this way, the perspective that choice contents are ever-present and designable may be. Jolting may admittedly have the additional down-side that it removes options from view more effectively than nudging, thus preventing the consideration of these options and so the moral development that such consideration could otherwise yield. On the other hand, if the options removed are immoral or otherwise undesirable, as they should be, the effect on moral development may be positive. By removing options more thoroughly, the effects of jolting may also be more resilient.

I conclude that the most common arguments against nudging are either unconvincing or come down to the worry that nudging can make nudgees instruments of another agent's will. I propose that this risk is associated with many different ways of influencing others and of shaping society, including some that we cannot do without. I propose further that this risk should be taken seriously in considering both nudging and jolting, but that it does in no way justify a general abstention from either tool.

7. Respect for Preferences

That preferences are indeterminate or context-dependent is one of the starting points for pro-nudgers. As I noted above, judgments too are indeterminate in the same way (indeed, a preference arguably is a form of judgment). This is why it is difficult to claim that any particular nudge promotes chooser wellbeing according to the choosers' own judgment. Even so, it is a central tenant of liberalism that individual preferences should be respected. This is presumably why Thaler and Sunstein defer the question of legitimate aims to chooser preference, despite the obvious problems associated with eliciting these preferences. Respect for preferences may be of independent importance, or it may be instrumental to the promotion of wellbeing. Either way, it is important, or so I will presume.

Concern with chooser preference is also a likely basis for arguments against jolting. Though options that are not accessible are often not considered and so not preferred, they *can* certainly be preferred. When they are, keeping them inaccessible conflicts with respect for preferences. Though I have rejected the freedom-preservation argument for nudges, most nudges do allow all or almost all choosers access to their preferred option at a cost that is at least not pro-

hibitive. Jolting, in contrast, makes options completely inaccessible, or prohibitively costly.

What does it mean that an inaccessible option is preferred? I explained above that it is morally irrelevant that some option would be chosen from a non-actual choice set, since this is true for (almost) all options.²² Accepting this, we could start elaborating on the degree to which some options are preferred over others over various possible choice sets. However, I fail to see why this would be relevant. Actual preferences should arguably be taken seriously as a matter of respect for persons and their choices. Preferences in non-actual choice situations warrant no such respect, because no one actually has them.

The fact that an option would have been preferred under non-actual circumstances should be distinguished from the fact that a non-present option is missed in the actual circumstances. Options that are not present in a choice cannot be chosen, of course, but they can still be preferred in the sense that we think, and perhaps state, that we would have chosen them if they were present.

We may also think that it would have been better to have access to certain options, even if we do not think we would have chosen them. This is essentially Berlin's argument, against Hobbes, for the importance of non-preferred options to the concept of (negative) liberty (2002[1969], 32). Whether or not we (think we) would have chosen it, we may call regret over the absence of some option *lack-of-option regret*. To illustrate: From the choice set {C, D, E, F}, the cafeteria costumer chooses C. I said this might be because she is attracted to the least unhealthy option, given that it is not too dull. This psychology is consistent with the customer regretting that A and B are not available, perhaps thinking (mistakenly, as it happens) that she would have chosen one of them, given the chance.

Lack-of-option regret is an actual state of mind, unlike preferences in non-actual circumstances. Lack of option regret is, or implies, a preference over choice sets. If the person experiencing lack-of-option regret could opt to have her choice set expanded to include the option she regrets not having, she would. To regret that an option is inaccessible is to prefer that the option set be expanded to include it. I propose that we should, *ceteris paribus*, respect preferences over choice sets, just as we should respect preferences over options in a choice set (indeed, sometimes choice sets *are* options).

Preferences over choice sets do not only come in the form of lack-of-option regret. As just noted, lack-of-option regret is a preference for expanding the current choice set. It is also possible to prefer a reduction of the current choice set. People often have this preference when they shop for utilities such as electricity, where the available options do not differ in the essential service provided, but this lack of relevant differences is obscured by how options are presented (with various creative marketing techniques).

We can have preferences over choice sets also for choices we are not currently facing. We can have preferences regarding specific choices, such as wishing there

²² There may be some options that are dispreferred in all possible choice sets in which they are members, but these options are very few.

will be a rich selection of sweaters in the store we plan on visiting tomorrow, or wishing that there will be no candy bars on sale in the supermarket the next time we shop there. We can also have more general preferences, such as a preference for facing fewer choices with large choice sets, because these choices tend to use up valuable time and effort. Importantly, we can have preferences over sets of choice sets that regard the aggregate content of those choices. For example, we may prefer not to have too many options to consume or purchase unhealthy foodstuffs any given day. We may think that it is fine to have an option to have a bun with the morning coffee, or an option to have a bun with the afternoon coffee, but not both these options. This is not a preference for any single choice set, but for the aggregate content of, in this case, two choice sets.

My brief investigation of preferences for inaccessible options has led me to note various possible preferences over choice sets. Lack-of-option regret can either consist in preferring an inaccessible option, or in preferring that some non-preferred option be accessible. The former preference is not always the more important. It is often more important to respect my preference for having access to a rich variety of options, than to respect my preference for there being some particular option that I would ideally like to choose.

I opened this section with granting the importance of respect for preference and the relevance of the argument against jolting that it can block preferred options. I end it by proposing that if respect for preferences is important, this includes preferences over choice sets, and so we must aggregate preferences for options and preferences for option sets in order to see how we can best respect preferences. Nudges and jolts of course typically target groups, and so the aggregate will include many preferences of each type. As preference-respecters, we will face difficult dilemmas, such as whether to respect some people's preference for making harmful options inaccessible, or respect other people's preference for keeping those options accessible, and perhaps for actually choosing them.

In conclusion: Choice content design sometimes operates just like choice context design in that it causes a chooser to prefer another option than she would have preferred in the absence of the design, without any negative experiences such as lack-of-option regret. While we should respect actual preferences, we have no reason to respect hypothetical preferences in non-actual choices. Our respect for actual preferences should not stop at preferences over options in choice sets, but should extend to preferences over choice sets, which come both in the form of preferences for some choice sets over others, and in the form of more aggregate properties of choice sets, such as the overall complexity and number of choices we must make.

8. Conclusion

In this article, I have defended both nudging and what I have called jolting. These are two different types of choice design. I have not been concerned to distinguish the two in detail, and it is not important for me that they can be distinguished, since I endorse both. On the other hand, the distinction is important for libertarian paternalists and others who endorse nudging but not jolting. Roughly speaking, nudging is design of the context in which a choice is made, while jolting is design of the content of the choice.

I have argued that nudging may sometimes have strong incentive effects, such as when a default must be set and the opt-out cost is quite high. Jolting has even stronger incentive effects, or may make options physically impossible to access. I have argued that strong incentive effects are not problematic as such, since they are often morally unavoidable in the sense that every choice must have some content and no particular content is morally neutral. I called this the argument from non-neutrality and explained how it is a modification of Thaler and Sunstein's argument from unavoidability.

Because of non-neutrality, different choice designs should be judged on their merits, irrespective of how they have come about. Littering public places with unmanned candy stores is not morally neutral, nor is abstaining from doing so. I granted that environments that have not been designed to influence us at all may be an exception to this general rule—perhaps such environments are morally neutral after all, or perhaps they have some special quality that overrides concerns with choice outcomes. Most choice environments, however, are not free of design in this strong sense.

I have considered three common arguments against nudging and how they may be relevant for jolting as well. I have found these arguments either unconvincing or dependent on the idea that nudging or jolting can make one agent the instrument of another agent's will. I have not addressed this very common concern, except by noting that it is not particular to nudging or jolting. I have explained how the arguments are no stronger against jolting than against nudging.²³

I have argued that neither the fact that jolting may limit liberty nor the fact that it may remove options that would otherwise have been preferred are reasons to reject jolting. I have investigated the different preferences we may have regarding our choices and I have endorsed the liberal position that, *ceteris paribus*, people should have what they prefer to have, whether their preferences concern options in single choice sets, the choice sets themselves, or aggregate properties of several choice sets. Respect for preferences, therefore, does not preclude either nudging or jolting, but rather makes them important tools for satisfying preferences.

²³ This strand of my argument could be taken as a reductio of the pro-nudging position, but this is not how I intend it.

In conclusion, the question is not whether nudging and jolting should be used, but what they should be used for, and when.

References

- Berlin, I. (2002[1969]), *Liberty*, New expanded edition of Four Essays on Liberty, Oxford: Oxford University Press.
- Beshears, J., J. J. Choi, D. Laibson and B. C. Madrian (2009), “The Importance of Default Options for Retirement Saving Outcomes: Evidence from the United States”, in: *Social Security Policy in a Changing Environment*, Chicago: University of Chicago Press, 167–195.
- Bornstein, R. F. (1989), “Exposure and Affect: Overview and Meta-Analysis of Research, 1968–1987”, *Psychological Bulletin* 106, 265–289.
- Bovens, L. (2009), “The Ethics of Nudge”, in: Grüne-Yanoff, T. and S. O. Hansson (eds.), *Preference Change: Approaches from Philosophy, Economics and Psychology*, Dordrecht: Springer, 207–219.
- (2012), “Real Nudge”, *European Journal of Risk Regulation* 1/2012, 43–46.
- Burgess, A. (2012), “Nudging Healthy Lifestyles: The UK Experiments with the Behavioral Alternative to Regulation and the Market”, *European Journal of Risk Regulation* 1/2012, 3–16.
- Buss, S. (2005), “Valuing Autonomy and Respecting Persons: Manipulation, Seduction, and the Basis of Moral Constraints”, *Ethics* 115, 195–235.
- Carter, I. and M. H. Kramer (2008), “How Changes in One’s Preferences Can Affect One’s Freedom (and How They Cannot): A Reply to Dowding and van Hees”, *Economics and Philosophy* 24, 81–96.
- Conly, S. (2013), *Against Autonomy: Justifying Coercive Paternalism*, Cambridge: Cambridge University Press.
- Dworkin, G. (1982), “Is More Choice Better Than Less?”, *Midwest Studies in Philosophy* 7, 47–61.
- (1988), “Behavior Control and Design”, in: *The Theory and Practice of Autonomy*, Cambridge: Cambridge University Press, 150–160.
- Grüne-Yanoff, T. (2012), “Old Wine in New Casks: Libertarian Paternalism Still Violates Liberal Principles”, *Social Choice & Welfare* 38, 635–645.
- Hausman, D. and B. Welch (2010), “Debate: To Nudge or Not to Nudge”, *Journal of Political Philosophy* 18, 123–136.
- Husak, D. (2013), “Penal Paternalism”, in: Coons, C. and M. Weber (eds.), *Paternalism: Theory and Practice*, Cambridge: Cambridge University Press, 39–55.
- Garnett, M. (2013), “Taking the Self out of Self-Rule”, *Ethical Theory and Moral Practice* 16, 21–33.
- Goodwin, T. (2012), “Why We Should Reject ‘Nudge’”, *Politics* 32, 85–92.
- Kahneman, D. (2011), *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux.
- Kuksov, D. and J. Miguel Villas-Boas (2009), “When More Alternatives Lead to Less Choice”, *Marketing Science* 29, 507–524.

- McIntyre, A. (2011), "Doctrine of Double Effect", in: Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition)*, URL: <http://plato.stanford.edu/archives/fall2011/entries/double-effect/> [08.08.2014]
- McNeil, B.J., S.G. Pauker, H.C. Sox Jr. and A. Tversky (1982) "On the Elicitation of Preferences for Alternative Therapies", *The New England Journal of Medicine* 306, 1259–1262.
- Pettit, P. (1997), *Republicanism: A Theory of Freedom and Government*, Oxford: Oxford University Press.
- (2008), "Republican Freedom: Three Axioms, Four Theorems", in: Laborde, C. and J. Maynor (eds.), *Republicanism and Political Theory*, Malden: Blackwell, 102–132.
- Saghai, Y. (2013), "Salvaging the Concept of Nudge", *Journal of Medical Ethics* 39, 487–93.
- Scanlon, T. M. (2008), *Moral Dimensions*, Cambridge: Harvard University Press.
- Shafir, E. (2013), *The Behavioral Foundations of Public Policy*, Princeton: Princeton University Press.
- Steiner, H. (1994), *An Essay on Rights*, Oxford: Blackwell.
- Sunstein, C. R. (1995), "Incompletely Theorized Agreements", *Harvard Law Review* 108, 1733–72.
- (2013), *Simpler: The Future of Government*, New York: Simon and Schuster.
- (2014), *Why Nudge? The Politics of Libertarian Paternalism*, New Haven: Yale University Press.
- and R. H. Thaler (2003), "Libertarian Paternalism Is Not an Oxymoron", *The University of Chicago Law Review* 70, 1159–1202.
- Thaler, R. H. and C. R. Sunstein (2009[2008]), *Nudge: Improving Decisions about Health, Wealth and Happiness*, 2d ed., New York: Penguin Books.
- Tversky, A. (1972), "Elimination by Aspects: A Theory of Choice", *Psychological Review* 79, 281–299.
- Vallgård, S. (2011), "Nudge—A New and Better Way to Improve Health?", *Health Policy* 104, 200–203.