

INTRODUCTION

A symposium on the behavioural limits of the state

Adam Oliver

Department of Social Policy, London School of Economics and Political Science, London, UK
Email: a.j.oliver@lse.ac.uk

(Received 13 April 2023; accepted 13 April 2023; first published online 29 May 2023)

Introduction

Behavioural public policy is a relatively new, but now established, subfield of public policy. It encompasses the theoretical and conceptual development, and associated application, of how behavioural science can inform the design of public policies, interventions and institutions. At its core, it is the study of how human behaviour reacts to context, and, as such, it is relevant to all conceivable aspects of public policy. Indeed, many public policy analysts work on issues that are foundational to behavioural public policy – for instance, in relation to social norms, identity, trust, etc. – without always realising it.

To date, behavioural public policy has been dominated by soft paternalistic policy frameworks that contend that people's contexts or environments can be altered in subtle ways that are informed by the findings of behavioural science, so that their automatic choices become more consistent with what they deliberately prefer. These frameworks are 'soft' because they do not explicitly force people to alter their behaviours; in principle, people can continue with their pre-existing behaviours if they so wish. The most famous manifestation of this approach is libertarian paternalism, applications of which are known as nudges.

Several criticisms can be levelled against soft paternalism, in relation to how effective these interventions are in addressing the fundamental challenges that societies face, on whether they are often actually informed by behavioural science (as opposed to standard assumptions of rationality), and on whether their inevitably often covert application really render them liberty preserving – i.e. that the hidden intentions that underpin many soft paternalistic interventions actually make them quite statist.

With normative questions on how behavioural science should, and should not, inform public policy in mind, I hosted a series of symposia on the appropriate 'behavioural limits of the state' through the offices of the Department of Social Policy and the Hayek Programme at the London School of Economics and Political Science from the Spring to the Autumn of 2022. A selection of the papers presented at those symposia form this special issue. In this Introduction, I offer a summary of the arguments

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

presented by the authors of those papers, arguments that lie at the core of behavioural public policy and the part that it plays and will play within broader social and public policy, as the field moves forward.

Summary of the arguments

A concern with the potential autonomy-eroding – or at the very least, insufficient autonomy-respecting – implications of soft paternalistic interventions was evident in the arguments made by several of the contributors to this special issue. Peter John (this issue), for instance, proposes that his nudge-plus framework, where nudges are accompanied with a detailed explanation of their intentions and motivations, are more ethically defensible than nudges alone, because the former allows the individual the opportunity to deliberate explicitly on whether they accept the behavioural intervention at the point of being exposed to it. In short, John concludes that it is better ethically to give some opportunity for reflection in the delivery of a nudge, an expression of support for free agency arguments that is shared by Ralph Hertwig (this issue).

Specifically, Hertwig argues that modern technologies employed by, for example, the food industry have become so sophisticated in their capacity to manipulate that they impose unacceptable harms on people over domains that would in the past have been widely viewed as within the private sphere of their lives. To tackle these harms, he calls for a comprehensive strategy that targets both supply-side interests and demand-side behaviours, and for the latter, he proposes a concept that he labels ‘self-nudging’ as an aspect of his conceptual framework called ‘boosting’, a framework that attempts to use the findings of behavioural science to improve individual agency. With self-nudging, Hertwig proposes that individuals are educated about the behavioural science basics, and then use that knowledge to become ‘citizen choice architects’; for example, to learn of the importance of salience, present bias and *status quo* effects so that they might place the biscuits at the back of their kitchen cupboards to reduce temptation (if they so wish), or so that they might set the default temperature a little lower on their home heating system. Hertwig suggests that with this approach, people remain autonomous over whether and how they should be ‘nudged’.

Cristina Bicchieri (this issue) focuses on ‘norm nudges’, which principally address externality concerns rather than having the paternalistic focus of the original nudge approach. Norm nudges steer individual behaviours by including some element of social comparison, providing information on what others do – e.g. 9 out of 10 people clean up their dog’s mess. Bicchieri contends that for these policies to have the best chance of working, the group used for comparison ought to be one with whom the target strongly identifies – e.g. for me, 9 out of 10 people in Lewisham clean up their dog’s mess. She further argues that in order to change a person’s behaviour through norm nudging, we need to change their social expectations, which requires the target and comparator group behaviours to be interdependent – i.e. the target group’s behaviour needs to be a function of what they believe others do or approve of. Integrating norm change interventions into public policy, therefore, requires one to carefully frame the message to appropriately target the specific interdependence that exists between people in real-world settings.

Returning to paternalistic nudges, in my article, I examine the arguments that are often made to suggest that people tend to make errors in their decisions and behaviours that they would ideally prefer to avoid (Oliver, this issue). These arguments focus on the four purported limitations of human reasoning – namely, limited imagination, willpower, objectivity and technical ability. I conclude that these purported limitations fail to provide an adequate justification for paternalistic interventions by policy makers for two interrelated reasons. First, despite common claims to the contrary, policy makers cannot really discern what individuals desire for their own lives, and second, so long as people are not harming others, they should be free of interference by third parties. I contend, perhaps to an even greater degree than do John and Hertwig, that paternalistic nudges impose unacceptable restrictions on individual autonomy and agency, and I call for a classical liberal – or, more specifically, Millian – framework for behavioural public policy.

Shaun Hargreaves Heap (this issue), in questioning whether people really have the stable, deliberative preferences that nudge advocates assume, also proposes a Millian behavioural public policy framework. He suggests that two insights, for the purposes of this argument, can be drawn from experiments in psychology and economics. First, people simply do not always behave in the way that rational choice theory predicts. They often employ a heuristic-driven process to reach their decisions and do not have an underlying fixed and stable preference ordering over many goods. The decision itself creates the apparent preference and not *vice versa*, and we cannot judge this to be irrational. Second, pro-social preferences – e.g. cooperative and even altruistic preferences – often exist, but can be heavily influenced by context. As such they are also unstable and can be crowded out in experimental settings, but, again, these context-specific preferences are not necessarily unreasonable. According to Hargreaves Heap, the lesson from these insights is that we should not attempt to steer people's behaviours towards some fixed and stable preference standard, and that we need to evaluate policy in other ways than preference satisfaction. As an alternative, he suggests that policies should be assessed over the degree to which they advance individual liberty in ways that enable them to gain a sense of autonomy. For this purpose, a negative sense of liberty (i.e. the absence of restriction) must be supplemented in two ways. One is the policy provision of some basic capabilities through, for example, an education system. The other comes from Mill's requirement that people should be free to do what they want but only so far as this does not cause harm to others (Mill, 1859/1969). The no-harm principle further requires a variety of familiar policies, albeit for less familiar reasons, like a health service and unemployment insurance. The principle also provides grounds for less familiar policies like a flat tax and a basic income.

Malte Dold and Paul Lewis (this issue) also discuss the importance of capabilities: specifically, what they term 'agentive' capabilities. They ground their argument by referring to Robert Sugden's (2018) focus on opportunities, in which he contends that the more options people have the better, irrespective of whether they have fixed and stable preferences. Dold and Lewis are similarly critical of the normative proposition, embedded in soft paternalism, that people ought to have fixed preferences, but they maintain that in addition to having opportunities, people want to be in control of the choice process. Manipulations and obfuscations by third parties,

informational complexity, and the possibility of being overwhelmed by too many choices, potentially erodes that control. Dold and Lewis maintain that people need opportunities *and* process freedom to protect their agentic capability, which can lead to perfect acceptable inconsistencies in choice. Like Hargreaves Heap, they contend that there ought to be a shift in emphasis in policy analysis from outcomes to rules and institutions that – similar to Hertwig’s boosts – improve people’s decision-making competences.

Sugden (this issue), in a related branch of his work, questions whether the anomalies that have been uncovered by behavioural economists undermine the validity of standard cost–benefit analysis (which is commonly used to appraise public sector investments), based as it is on stated willingness-to-pay exercises. He sketches the outline of a cost–benefit methodology that is compatible with inconsistent preferences, in accordance with a well-functioning market as a network of opportunities for mutually beneficial transactions. Sugden illustrates his argument with the example of fuel economy standards, where he shows that cost–benefit analysis has tended to classify reduced fuel consumption in and of itself as a benefit. He argues that this ought not count as a benefit in a formal analysis; rather it is the consequence of individuals engaging in a mutually beneficial exchange where one party (the consumer) pays another party (the car retailer) a premium price for a good that he or she desires (in his case, a car that consumes less petrol). The only external benefits that ought to be incorporated into the cost–benefit ratio in this scenario, according to Sugden, are the reduced exhaust emissions that would otherwise harm public health and the environment.

Sugden implicitly places less faith in government intervention as a means of securing optimal societal benefits than is justified (at face value) by many cost–benefit analyses. Michael Romero and Virgil Storr (this issue) also believe that we should look less towards policy makers to make decisions for us, but for different reasons. They argue that in-group preferences and biases and out-group prejudice, which they define as a form of parochialism, can cause individuals to evaluate policy based on how it affects the groups they belong to rather than its impact on society at large, or even on their ideological priors. To counter the possibility of this parochialism influencing policy maker decision-making, Romero and Storr call for limited government and an open society, but mirroring the arguments made by Hargreaves Heap and Dold and Lewis, they emphasise the need to foster an institutional environment that encourages people to interact with others as diverse equals.

Liya Palagashvili and Nathan Goodman (this issue) also look at how institutions might be designed so as to improve the performance of boundedly rational actors, in a consideration of how behavioural public policy might apply to the field of criminology. They contend that unless one is careful, the institutional context can interact with individual decision-making heuristics to cause harms to others. They highlight this argument by claiming that the militarisation of the police force in the United States has primed many officers to see people as enemies – to return to Romero and Storr, as outgroups – that must be controlled rather than as citizens that they must protect and serve. They further suggest that since police officers are trained to deal with violent situations but spend much of their time dealing with innocuous events, there is a mismatch between the officers’ perceptions and reality. As a

solution, Palagashvili and Goodman call for decentralised ‘polycentric’ systems of policing, so that officers are better matched, and closer to, the predominant challenges that they face in their daily working lives. This, they suggest, would produce a more mutually beneficial relationship with the communities that the police force serves.

Conclusion

As this summary, and the articles that follow, hopefully demonstrate, the normative issues pertaining to the future development of the field of behavioural public policy are as rich as any of those found in the social sciences. Several of the articles challenge the welfarist and consequentialist approaches that have dominated rational choice theory, behavioural economics and behavioural public policy to date, and highlight instead the importance of rules and institutions. These articles in essence present contractarian arguments that favour individual autonomy, while at the same time recognise that rules and institutions are required to contain egoistic tendencies and to nurture pro-social motivations. Of course, for welfarists, and perhaps others, the arguments presented are in many cases contentious, but no responsible public policy can be implemented without first thinking deeply about the appropriate limits and design of state action.

References

- Mill, J. S. (1859/1969), ‘On Liberty’, in: *On Liberty and Utilitarianism*. Oxford: Oxford University Press, 1–110.
- Sugden, R. (2018), *The Community of Advantage: A Behavioural Defence of the Liberal Tradition of Economics*. Oxford: Oxford University Press.