

Three challenges for behavioural science and policy: the empirical, the normative and the political

ROBERT LEPENIES

Helmholtz Centre for Environmental Research – UFZ, Leipzig, Germany

KATHRYN MACKAY*

Department of Politics, Philosophy and Religion, Lancaster University, Lancaster, UK

MUIREANN QUIGLEY

Birmingham Law School, University of Birmingham, Birmingham, UK

Abstract: In a ‘post-truth’ era in which personality and opinion trump evidence and reason, the need for frankness in debates about the use and boundaries of science and policy is high. We welcome the reflective and nuanced approach to behavioural science in policy-making in Sanders, Snijders and Hallsworth’s (2018) piece. Despite our support for the approach in this paper, we suggest that there are deeper issues than are currently acknowledged. Our critique tackles three issues: the empirical, the normative and the political. In the first section, we examine what counts as ‘behavioural’ and how this label is used to legitimate a range of policy activities. We then look at randomised controlled trials in the next section, highlighting the extra-scientific dimensions of the empirical ‘What Works’ revolution. Finally, we question some ontological assumptions that drive empirical research and its translation into policy, asking where the collective is to be found in behavioural public policy.

Submitted 17 January 2018; accepted 13 March 2018

Introduction

As part of the Brexit ‘Leave’ campaign, Michael Gove famously claimed that “people in this country have had enough of experts” (Mance, 2016). Contrary to what one might infer from recent reporting on the role of expertise and Brexit, public trust in science and scientists is at a near all-time high (British Science Association, 2014; IPSOS Mori, 2017). Nevertheless, the political

* Correspondence to: Department of Politics, Philosophy and Religion, Lancaster University, Lancaster LA1 4YQ, UK. Email: k.mackay1@lancaster.ac.uk

climate has been deeply impacted by ‘fake news’, with some arguing we have arrived in a ‘post-truth’ era in which personality and opinion trump evidence and reason. This makes the need for frankness in debates about the use and boundaries of science and policy all the more essential. We therefore welcome the more reflective and nuanced approach to the use of the behavioural science in policy-making that we detect in Sanders, Snijders and Hallsworth’s (2018) piece and other recent work by the Behavioural Insights Team (BIT; e.g., Hallsworth & Sanders, 2016). In their account of where we are now with behavioural science and policy, Sanders, Snijders and Hallsworth (2018) comment that “there is much for behavioural science to be pleased with.” Yet despite our support for the approach taken in their paper, we suggest that there are fewer reasons to be pleased than currently acknowledged. Our critique tackles three main issues: the empirical, the normative and the political.

To these ends, in the first section, we examine the twin issue of what counts as ‘behavioural’ and how this label is being used to legitimate a range of policy activities. We then look specifically at randomised controlled trials (RCTs) in the next section, highlighting the extra-scientific dimensions of this empirical ‘What Works’ revolution (What Works Network, 2014). Following this, we also question some of the ontological assumptions that drive empirical research and its translation into policy; in particular, we ask where the collective is to be found in behavioural public policy.

Behavioural interventions, behavioural science and legitimacy?

The term ‘behavioural sciences’ is often used in policy debates without due attention to what this is actually referring to. For a representative example, see the contributions on health nudges in Cohen, Lynch and Robertson (2016), in which solutions derived from behavioural science are discussed – without once providing a definition of behavioural sciences. This matters because different academic fields have a different understanding of what interventions fall under its remit, and it becomes unclear to what the term refers or how exactly it links to fields like behavioural economics or cognitive psychology (and how, for instance, it is separate from other fields such as evolutionary psychology).

Currently, the use of the term ‘behavioural sciences’ fulfils an interesting justificatory role at BIT and similar organisations. Proponents of behavioural policy draw on the scientific authority of various fields of study on behaviour to *justify* policy interventions (Lepenies & Malecka, 2015; Lepenies & Malecka, [Forthcoming](#)). Yet this scientific authority itself is rather amorphous (Malecka & Nagatsu, 2017). As a result, there is a risk of claiming more for

behavioural science than would hitherto have been considered to count. This tendency is evident in Sanders, Snijders and Hallsworth's (2018) current contribution to debate. Their piece, we suggest, reveals a more general and prevalent temptation to look at everything policy related through a 'behavioural' lens (e.g., Hallsworth, 2016). The consequence of such framing is that substantial swathes of policy end up being viewed as if they are *centrally concerned* with behaviour. While this may be true in a trivial sense – for instance, we might say that all regulation is about behaviour in the sense that it aims to alter the behaviour of some actor(s) or other (Quigley, 2014) – the breadth of such a framing is difficult to legitimise.

There are two main issues here. First, the authors are overly permissive in what counts as a behavioural intervention, resulting in what we would like to call 'behavioural policy overreach'. Second, and something that we tackle in the third section of this paper, is a problem that arises with the use of RCTs. In essence, there is a tendency to claim an intervention to be a 'behavioural' policy tool in virtue of the fact that it has been tested via RCT (even for traditional regulatory tools such as a taxes or mandates). Combined, these two aspects become self-legitimising; that is, they support policy-makers' claims to engage in all types of policy interventions on *behavioural* grounds, which are then claimed to be more 'scientific'.

An example of such behavioural policy overreach can be seen where Sanders, Snijders and Hallsworth (2018) describe governmental regulations that ban or put limits on food ingredients like sugar or trans-fats. Such a regulation, governing the addition of ingredients in food, is not a behavioural stimulus. It is traditional command-and-control regulation, where certain constraints are *imposed* on those regulated; in this case, manufacturers. Relatedly, the authors argue that, sometimes, behavioural analyses say that the best intervention is not to change behaviour, but to *mitigate the consequences* of the behaviour. Thus, such analyses encompassed within 'behavioural' policies do not seek to change people's consumption habits, but instead change the outcomes that accrue. But again, this ought not to be construed or presented as a behavioural intervention. It is not seeking to change the behaviour in question, but rather its corollaries. In a sugar reduction policy, such as the one the UK implemented in April of 2018, the goal of the policy is to have companies change the formulation of their products to lower the amount of sugar people consume without changing their behaviour. This is simply old-fashioned regulation, with higher taxation as a penalty to those who do not comply (and is one that appears to work) (Hallsworth, 2016).

This temptation to claim standard forms of regulation or legislation as part of 'the behavioural sciences' is distorting. Such regulation can be clearly distinguished. Moreover, it is good and proper to do so. The European

Commission's Joint Research Centre (JRC) has suggested distinguishing between behavioural interventions by classifying them into behaviourally informed, behaviourally aligned and behaviourally tested initiatives (Lourenço, *et al.*, 2016). However, simply because a policy is discussed by a behavioural insights team does not render it 'behavioural' in any of these proposed classifications. Without clear demarcations of what 'behavioural interventions' can and cannot achieve, the danger is that behavioural policy is mistakenly identified with 'whatever behavioural policy units do'. This harbours political risk for evidence-based policy in government and for the scientific evidence base it aims or claims to draw upon. Even the most genuine attempt to take a more evidence-based approach to policy cannot apply scientific rigour to all aspects of its operations. Nudge units are also political actors, part of whose operations simply defy evidence-based principles. To take just one example, note that the process of 'issue selection', by which new behavioural interventions or new fields of engagement are chosen, likewise is not informed by evidence only. Our claim is not that nudge units *always* engage in such justifications of policies, but that care needs to be taken, with the emphasis put on 'evidence', what counts and whether it answers the questions we need it to, as we discuss below.

Explicitly demarcating and acknowledging the limits of the behavioural sciences – both empirically and normatively – may reduce the scope of the application of 'behavioural' tools to policy-making, as well as the political influence of such approaches, possibly beyond a point that the authors and other proponents would find desirable. Yet it might be a long-term strategy that is in their interest. Claiming more for (behavioural) science than is due puts the credibility and legitimacy of behavioural approaches at risk. Conversely, there is little risk (beyond perhaps the purely political) and much to be gained by saying, "Here are the limits of behavioural science as it applies to policy, and beyond this it has nothing to say about how to form good policy."

A bundle of assumptions

The justification supporting many 'behavioural' policy proposals is rooted in a notion that scientifically grounded results are undeniably true, accurate and authoritative. In particular, much justificatory weight within behavioural policy practice is placed upon RCTs. We therefore examine some issues that attend these, before turning to consider the individualistic focus of much of behavioural public policy.

RCTs, evidence and politics

BIT is firmly committed to the use of RCTs to test policy interventions. This is clear from the 300+ such trials that they have conducted in recent years

(Sanders *et al.*, 2018). Much of what we write here will not be news to practitioners (see Glennerster, 2013), and we are not critical of the use of RCTs per se. Our concern is that the use of RCTs in the behavioural policy arena is outpacing critical reflection about their uses and the distinct epistemic and normative presumptions of the ‘policy culture’ in which they are drawn upon.

Proponents and critics alike should welcome the advanced reflections on the methodological limits of RCTs in the work of those within BIT (e.g., Hallsworth, 2016; Hallsworth & Sanders, 2016; Sanders *et al.*, 2018). Nevertheless, we push for further contextual engagement with these here. There are issues not only about the (types of) *answers* that RCTs can provide, but also about the (types of) *questions* they are suited to ask in the first place. The type of behavioural public policy practiced by BIT comes with an implicit (only sometimes explicit) theory of evidence. There is a tendency to assume that RCTs can establish ‘what works’. However, as a number of commentators have argued elsewhere, such trials are not necessarily suited to answering the ‘what works’ question when applied to social policy unless a long list of qualifications is added and addressed (e.g., see Reiss, 2007, 2014; Deaton, 2009; Cartwright, 2010; Favereau, 2016; Deaton & Cartwright, 2017). As Cartwright and Hardie (2012) have pointed out, RCTs generally establish not ‘what works’, but only that something worked, at some place, at that time. Cartwright (2010) goes a step further and argues that all that RCTs can do in terms of policy predictions is to predict the efficacy of similar RCTs with the same target population. They are good at identifying causal connections between interventions on a target and outcomes under very *specific* conditions and circumstances. Only with a large host of additional assumptions can the external validity of RCTs be guaranteed and the promise of social effectiveness be achieved. This means that their use and application in the area of social policy interventions can be problematic, especially where a general roll-out of policy is intended after limited trials. This is all known to practitioners. But beyond questions of effectiveness lurk other distinct issues with RCTs in social policy that can only be understood in a broader context (Jung *et al.*, 2014).

A look at more matured debates about the use of RCTs in clinical contexts is instructive here (Ashcroft, 2004) and shows that there are different ‘cultures’ of how much we can learn from experimental evidence in different settings. Favereau (2016), for example, notes that RCTs in medicine have historically included patient safety and effectiveness protocols, whereas in social policy RCTs (or randomised controlled experiments), only effectiveness is tested. Safety – social or psychological – is not included. This, of course, is not specific to RCTs as a method, but depends upon the disciplinary context of

how they are understood. Therefore, more attention should be paid to the specific implications for the culture of policy-making of widespread RCT usage.

Behavioural science researchers have been attentive to the ethics of experimental trials (Glennerster, 2013), but it is the political ramifications of policy-by-RCT and their implications specifically within the culture of policy-making that are missing. There are issues with the (implicit) theories of desired political change that the widespread use of RCTs engenders. At the level of behavioural policy in practice, RCTs are used in the context of a project of ‘radical incrementalism’. As outlined by Halpern, this “is the idea that dramatic improvements can be achieved, and are more likely to be achieved, by systematically testing small variations in everything we do, rather than through dramatic leaps into the dark” (2015, p. 291). Added to this, a look at the use of RCTs in development policy shows another policy culture in which RCTs are used. Banerjee and Duflo (2009) speak here of a “quiet revolution” through RCTs, as they argue that political constraints “make it difficult to find big solutions to big problems” (for critiques, see Reddy, 2012; Berndt & Boeckler, 2016). Proponents of behavioural policy may agree with these normative theories of political change or they may disagree. However, it means that policy-by-RCT is neither ideologically neutral nor purely evidence based; it depends on the policy culture in which it is utilised. Sanders and colleagues (2018) write that after the successes of a maturing BIT, it is possible to focus more on longer-terms projects and less on politically expedient ‘quick wins’. We welcome this as an opportunity for BIT and similar units to also think critically about the policy culture that will mature with this.

Whither the collective?

Our penultimate concern about the description of behavioural advances in Sander, Snijders and Hallsworth’s (2018) article is about the individualistic focus that behavioural policy-making seems to take. While we are encouraged by their consideration of upstream regulatory change, as they mention in their discussion of corporate fossil fuel use, the authors tend to focus on individuals, even when thinking about groups. For example, in the section about social diffusion, they say that the behavioural sciences have not focused enough on understanding how to harness peer-to-peer networks in order to influence behaviour. We do not dispute the potential (in terms of policy/intervention effectiveness) that drawing on social norms and dynamics represents. We do, however, see two issues that bear further reflection.

First, future work in this area would do well to think beyond the individual. Frequently, in considering social groups, the focus is on how *one individual* may impact the behaviour of *another individual*. Even organisations are just

a sum of individuals. Again, there is a trivial sense in which this is true; of course groups are composed of the individuals that make them up. However, groups are more than this in important ways. For instance, Sanders *et al.* (2018) do not make mention of any *collective* notion of identity and agency. Yet, these may be key to understanding *why* the behaviour of groups can be different from the behaviour that each individual within the group would display if they were separate from the group (a negative example is of rioting and a positive example is of standing up to a bully; see the various discussions in Tajfel, 1982). Solidarity, community, compassion and a number of other values are in play with regards to group agency, but do not seem to ordinarily be taken into account or examined in the behavioural policy arena.

Second, there has been very little critical reflection upon the *permissibility* of drawing on social dynamics, either by behavioural policy practitioners or theorists. Studies focus on the *effectiveness* of exploiting social norms without explicit consideration of the potential damage to relationships within communities (e.g., Kettle *et al.*, 2016). The difficulty with this is that such approaches may not discriminate between instruments that address citizens as reasonable agents and those that treat them merely as social beings/members of communities responding to cues. There is, therefore, work to be done on the (normative) limits of this kind of policy-making. As such, there are a number of significant, yet unanswered questions. These include how behavioural policy-makers conceive of the citizen in social relationships; how (or whether) they account for family-based decision-making or group alienation from a system that is attempting to organise society; and how they account for behaviours that appear to be undesirable, irrational or unwise from their perspective, but that are eminently desirable, rational and wise from a different point of view without falling back on the ‘cognitive bias’ notion (e.g., see work on ‘shortened horizons’ by Warin *et al.*, 2015; or that on minority group women who decide to have children at an earlier age by Geronimus, 1987, 1996).

If behavioural policy-makers claim to understand human behaviour to the point that they can reliably predict it, and manipulate it, then they need to include different conceptualisations about groups. Much of our lives are led within larger or smaller groups, and there is reason to doubt that we can really understand what individuals do, and why, without considering their social/familial group identities. This is something that could be aided by deeper consideration of different psychological theories regarding how individuals and groups act, as well as different normative theories about what values and principles ought to be given precedence in policy-making.

Concluding thoughts

As we noted at the beginning of this commentary, we welcome the more reflective and longer-term thinking evident in the article by Sanders and colleagues (2018). When charting the development of behavioural public policy, there is, as they say, much to be pleased with. However, we end on a note of caution. The authors say that the principle of having “a positive social impact,” inspired by Thaler’s ‘nudge for good’ (Thaler, 2015), is at the core of BIT practices today. Behavioural policy-makers will, of necessity, and by favouring some interventions over others, take a stand on empirical, normative and political issues. This includes giving implicit answers about what counts as ‘positive social impact’ and what is ‘good’. Given this, we would encourage further work and reflection on the assumptions and choices in order to make them more transparent and explicit and to foster crucial debate regarding these matters.

References

- Ashcroft, R. E. (2004), ‘Current epistemological problems in evidence based medicine’, *Journal of Medical Ethics*, 30(2): 131–135.
- Banerjee, A. V. and E. Duflo (2011), *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, USA: Public Affairs.
- Berndt, C. and M. Boeckler (2016), ‘Behave, global south! Economics, experiments, evidence: editorial’, *Geoforum*, 70: 22–24.
- British Science Association (2014), ‘Public Attitudes to Science Survey’, Available at: <https://www.britishsociety.org/public-attitudes-to-science-survey>
- Cartwright, N. (2010), ‘What are randomised controlled trials good for?’ *Philosophical Studies*, 147: 59–70.
- Cartwright, N. and J. Hardie (2012), *Evidence-Based Policy: A Practical Guide to Doing it Better*, Oxford, UK: Oxford University Press.
- Cohen, I. G., H. F. Lynch and C.T. Robertson Eds. (2016), *Nudging Health: Health Law and Behavioural Economics*, Baltimore, USA: Johns Hopkins University Press.
- Deaton, A. (2009), ‘Instruments of Development: Randomization in The Tropics, and The Search for The Elusive Keys to Economic Development’, *National Bureau of Economic Research Working Paper No. 14690*, 1–54.
- Deaton, A. and N. Cartwright (2017), ‘Understanding and Misunderstanding Randomized Controlled Trials’ (Revised version), *NBER Working Paper, No. 22595*, the National Bureau of Economic Research.
- Favereau, J. (2016), ‘On the analogy between field experiments in economics and clinical trials in medicine’, *Journal of Economic Methodology*, 23(2): 203–222.
- Geronimus, A. T. (1987), ‘On Teenage Childbearing and Neonatal Mortality in the United States’, *Population and Development Review*, 13(2): 245–279.
- Geronimus, A. T. (1996), ‘What teen mothers know’, *Human Nature*, 7(4): 323–352.
- Glennester, R. (2013), *Running Randomized Evaluations: A Practical Guide*, New Haven, USA: Princeton University Press.
- Hallsworth, M. (2016), ‘The softdrinks levy is working before it has even been applied’, *Behavioural Insights Team Blog*, November 11. Available at: <http://www.behaviouralinsights.co.uk/unategorized/the-soft-drinks-levy-is-working-before-it-has-even-been-applied/>

- Hallsworth, M. (2016), 'Seven Ways of Applying Behavioral Science to Health Policy', in C. T. Robertson, I. Glenn Cohen, and H. Fernandez Lynch (eds), *Nudging Health: Health Law and Behavioral Economics*, Johns Hopkins University Press, pp. 40–51.
- Hallsworth, M. and M. Sanders (2016), 'Nudge: Recent developments in behavioural science and public policy', in F. Spotswood, Ed., *Beyond Behaviour Change: Key Issues, Interdisciplinary Approaches, and Future Directions*, Bristol, UK: Policy Press.
- Halpern, D. and D. Mason (2015), 'Radical Incrementalism', *Evaluation*, 21(2): 143–149.
- IPSOS Mori (2017), 'Trust in Professions: Long-term trends', 30 November. Available at: <https://www.ipsos.com/ipsos-mori/en-uk/trust-professions?view=wide>
- Jung, A., R. L. Korinek and H. Straßheim (2014), 'Embedded expertise: a conceptual framework for reconstructing knowledge orders, their transformation and local specificities', *Innovation: The European Journal of Social Science Research*, 27(4): 398–419.
- Kettle, S., M. A. Hernandez-Hernandez, S. Ruda and M. Sanders (2016), *Behavioral Interventions in Tax Compliance: Evidence from Guatemala*, Policy Research working paper N° WPS 7690; Impact Evaluation series. Washington, D.C.: World Bank Group.
- Lepenies, R. and M. Małecka (Forthcoming), 'Behaviour Change: Extralegal, apolitical, scientific?', in S. Beck and H. Strassheim (eds), *Handbook of Behavioural Change and Public Policy*, UK: Edward Elgar Publishing.
- Lepenies, R. and M. Małecka (2015), 'The Institutional Consequences of Nudging – Nudges, Politics, and the Law', *Review of Philosophy and Psychology*, 6(3): 427–437.
- Lourenço, J. S., E. Ciriolo and S. R. Almeida and X. Troussard, Joint Research Centre (2016), *Behavioural Insights Applied to Policy: European Report 2016* Brussels: European Union, http://publications.jrc.ec.europa.eu/repository/bitstream/JRC100146/kjna27726enn_new.pdf
- Małecka, M. and M. Nagatsu (2017), 'How behavioural research has informed consumer law: The many faces of behavioural research' in K. Purnhagen, A. Sibony and H. Micklitz (eds.), *Research Handbook on Methods in Consumer Law*, UK: Edward Elgar Publishing.
- Mance, H. (2016), 'Britain has had enough of experts, says Gove', *Financial Times*, 3 June. Available at: <https://www.ft.com/content/3be49734-29cb-11e6-83e4-abc22d5d108c>
- Quigley, M. (2014), 'Are Health Nudges Coercive?' *Monash Bioethics Review*, 32(1–2): 141–158.
- Reddy, S. G. (2012), 'Randomise this! On poor economics', *Review of Agrarian Studies*, 2(2): 60–73.
- Reiss, J. (2007), *Error in Economics: Toward a More Evidence-Based Methodology*, Oxford, UK: Routledge Publishing.
- Reiss, J. (2014), 'Struggling Over the Soul of Economics: Objectivity Versus Expertise', in C. Martini, M. Boumans, Eds., *Experts and Consensus in Social Science. Ethical Economy*, Springer International Publishing, pp. 131–152.
- Sanders, M., V. Snijders and M. Hallsworth (2018), 'Behavioural science and policy: where are we now and where are we going?' *Behavioural Public Policy*, 2(2): 144–167.
- Tajfel, H., Ed. (1982), *Social Identity and Intergroup Relations*, Cambridge, UK: Cambridge University Press.
- Thaler, R. H. (2015), 'The Power of Nudges, For Good and Bad', *The Upshot, New York Times*, 31 October. Available at: https://www.nytimes.com/2015/11/01/upshot/the-power-of-nudges-for-good-and-bad.html?_r=0
- Warin, M., T. Zivkovic, V. Moore, P. R. Ward and M. Jones (2015), 'Short horizons and obesity futures: Disjunctures between public health interventions and everyday temporalities', *Social Science and Medicine*, 128, 309–315.
- What Works Network (2014), 'What works? Evidence for decision makers', 25th November. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/378038/What_works_evidence_for_decision_makers.pdf