

Behavioural science and policy: where are we now and where are we going?

MICHAEL SANDERS*

Behavioural Insights Team, London, UK and University of Oxford, Blavatnik School of Government, Oxford, UK

VEERLE SNIJDERS

Behavioural Insights Team, London, UK

MICHAEL HALLSWORTH

Behavioural Insights Team, London, UK and Imperial College London, London, UK

Abstract: The use of behavioural sciences in government has expanded and matured in the last decade. Since the Behavioural Insights Team (BIT) has been part of this movement, we sketch out the history of the team and the current state of behavioural public policy, recognising that other works have already told this story in detail. We then set out two clusters of issues that have emerged from our work at BIT. The first cluster concerns current challenges facing behavioural public policy: the long-term effects of interventions; repeated exposure effects; problems with proxy measures; spillovers and general equilibrium effects and unintended consequences; cultural variation; ‘reverse impact’; and the replication crisis. The second cluster concerns opportunities: influencing the behaviour of government itself; scaling interventions; social diffusion; nudging organisations; and dealing with thorny problems. We conclude that the field will need to address these challenges and take these opportunities in order to realise the full potential of behavioural public policy.

Submitted 25 June 2017; accepted 13 March 2018

Introduction

The year 2018 marks nine years since Cass Sunstein entered the US federal government as Administrator of the White House Office of Information and Regulatory Affairs (OIRA) and eight years since the creation of our employers, the UK Behavioural Insights Team (BIT). This period has seen many other

* Correspondence to: Behavioural Insights Team, 4 Matthew Parker Street, London SW1H 9NP, UK. Email: michael.sanders@bi.team

attempts to apply behavioural science to government; various others, notably Sunstein (2014), Thaler (2015) and Halpern (2015), document and reflect on these developments. We do not attempt to cover the same ground here.

Instead, we first briefly summarise BIT's origins, before reviewing and assessing the current state of behavioural science in policy. This is followed by a more in-depth discussion of areas that have not been central to the application of behavioural science to policy so far. In our view, these represent either the most substantive critiques or the most interesting avenues for future investigation. We conclude by suggesting what changes would be needed to ensure these areas are explored in the future.

The origins of the Behavioural Insights Team

We start close to home. In 2010, BIT became one of the first government institutions (if not the first) dedicated to applying behavioural science to policy and public administration. Set up as part of the UK Prime Minister's Office and the Cabinet Office, the seven members were set three objectives to achieve in order to avoid triggering a 'sunset clause' that would see the team shut down on its two-year anniversary. These were: (1) transform at least two major areas of policy; (2) spread an understanding of behavioural approaches across Whitehall; and (3) achieve at least a tenfold return on cost.

In addition to these objectives, BIT also developed two main guiding principles, which are still core to its practices today. The first principle was to have a positive social impact, inspired by Richard Thaler's mantra of 'nudge for good'. This principle has not only influenced BIT's choice of projects, it has also meant making our findings publicly available in order to spread the use of behavioural approaches as widely as possible. The second principle was to robustly evaluate the impact of its interventions. Often, this principle was realised by promoting the use of randomised controlled trials (RCTs) wherever possible, including in public administration contexts where they had previously been uncommon. We summarised this principle as 'test, learn, adapt' (Haynes *et al.*, 2012).

We also adopted some strategies to address the specific challenges we faced in our early days. One was that we would focus initially on translating the best-evidenced interventions from the behavioural science literature in order to provide a proof of concept and some 'quick wins'. This would have the additional benefit of addressing concerns that effects generated in laboratory studies would not translate into real-world contexts (Institute for Government and Cabinet Office, 2010). Secondly, it was decided to focus on revenue-producing or money-saving projects. Not only would such projects address Objective 3 and help build the case to maintain the team, but also

they could often draw on practical advantages such as reliable data sources and established communication mechanisms. The combination of our initial focus on small changes to existing processes, the connection to Thaler and Sunstein's book (2008) and the generous support of Thaler himself led to the team attracting the nickname of 'the Nudge Unit'.

Where are we now?

In order to meet increasing demand in the UK and abroad, BIT 'spun out' of government in 2014 and became a social purpose company. The organisation is still partially owned by the UK government, but is now shared with Nesta (the innovation charity) and the employees themselves – with each of these partners owning a third of the company. The team now has over 120 staff and offices in London, Manchester, New York, Sydney, Singapore and Wellington. It has strategic partnerships with several universities, including Harvard, University College London and Imperial College London. Its 2015–2016 annual update contained the results of 46 individual randomised trials (BIT, 2016). Despite these changes, BIT has maintained its focus on the guiding principles of social impact and robust evaluation.

At the same time, other UK government departments have established their own behavioural science teams, including HM Revenue and Customs (who have a team twice the size of BIT's original Cabinet Office incarnation), Public Health England, the Department of Work and Pensions and the Department of Education. This expansion has not been limited to the UK, however: many countries now explicitly use behavioural science in policy settings. The Organisation for Economic Co-operation and Development (OECD)'s report on behavioural insights (OECD, 2017) contains 150 case studies of such applications gathered from around the world – and this list is illustrative rather than exhaustive. Dedicated teams have been set up in the US federal government, the United Nations, the World Bank and in more than a dozen other countries and organisations around the world. Many reports on the application of behavioural science worldwide have been written (e.g. World Bank, 2015). Other jurisdictions have made use of behavioural insights through non-governmental organisations and consultancies that work mainly or exclusively in the area (Whitehead *et al.*, 2014). The scale and duration of activity led the OECD to conclude that "Behavioural insights can no longer be seen as a fashionable short-term foray by public bodies. They have taken root in many ways across many countries around the world and across a wide range of sectors and policy areas" (OECD, 2017, p. 13).

Much work has also gone into the development and dissemination of ideas to support (or question) these activities. Dedicated journals, such as this one,

have been created to consider the relationship between behavioural science and policy. Several edited volumes on the topic are either in circulation or in press (Oliver, 2013; Shafir, 2013; Pykett *et al.*, 2017). Many universities have set up centres, groups and networks dedicated to the application of behavioural sciences to public sector issues (and some to critique attempts to do so). An increasing number of international conferences explicitly bring together academics and practitioners interested in the topic – and there appears to be the demand to sustain them. The UK government now requires policy-makers to have received training in principles from behavioural science.

There is much for proponents of behavioural science to be pleased with here, but it is difficult to argue that it constitutes a revolution. There have been behavioural insights teams that have failed to get off the ground or that have been launched and failed to make a meaningful contribution – whether through contingent factors or deficiencies in ability. While behavioural science is much more widely used than it was, it has yet to sit alongside economics as a discipline dominant in the thinking of policy-makers. For example, BIT has conducted over 400 RCTs in the last seven years on policy issues including charitable fundraising, changing general practitioners' prescribing behaviour, getting people back into work and stopping people from reoffending. However, this obviously only represents a tiny fraction of total public sector activity. Moreover, as discussed below, there is a danger that behavioural science is seen to offer merely technocratic tweaks, rather than the more wide-ranging reassessment of public administration that could be possible.

The 'replication crisis' currently gripping psychology (and other sciences) is also creating wider and more profound consequences for behavioural science and policy. A combination of questionable research practices and small sample sizes, particularly in laboratory experimental research, means that many previously accepted findings are now being called into question. The crisis should not be dismissed as of merely academic interest, since several of these findings are ones that have been – or could be – applied to policy problems. What the global community of behavioural scientists does next will determine whether policy-makers will continue to see behavioural science as a reliable source of policy ideas and approaches.

Where do we go from here?

Below we present two clusters of ideas. The first deals with the complications and challenges that face behavioural public policy: the long-term effects of interventions; repeated exposure effects; problems with proxy measures; spillovers and general equilibrium effects and unintended consequences; cultural

variation; ‘reverse impact’; and the replication crisis. The second cluster concerns opportunities: influencing the behaviour of government itself; scaling interventions; social diffusion; nudging organisations; and dealing with thorny problems. We recognise that there may be overlaps between these topics.

Cluster 1: Complications and challenges

Long-term effects

One common concern is that we lack evidence that effects on behaviour endure for a substantial period. Despite many field experiments testing the application of psychology or behavioural economic interventions to practical problems (a good summary of which can be found in DellaVigna, 2009), relatively little is known about the long-term effects of many such interventions (Frey & Rogers, 2014). BIT has experienced this problem because it has been required to range widely across many policy areas, which has sometimes limited the incentives to return to an intervention in order to assess its long-term consequences. We are beginning to address this omission in some areas. We know, for example, that non-compliant Guatemala taxpayers who received either a social norm or an ‘oversight message’ were significantly more likely to comply 12 months later, despite not receiving another letter in the interim (Kettle *et al.*, 2016). We are also engaged in a long-term evaluation of the economic impact of receiving a growth voucher, one of the largest RCTs of its kind.¹

One underlying issue is that public officials and academics (particularly junior scholars) are rarely incentivised to choose studies where the main outcome measure will only be reported far in the future.² Obviously, a longer discussion is needed to deal with the various aspects of this problem, so we only sketch some ideas here. One is that public officials applying behavioural science should plan to achieve a balance between longer-term outputs and the short-term outputs that they often need to justify continued support and funding. Another possibility is to anticipate that personnel and attention will shift as time moves on and mitigate the consequences. For example, those running trials could be required to leave them in a state that allows

¹ Growth vouchers were matched with government funding of up to £2000 for small businesses to access expert business advice and support. More information is available at: <https://www.gov.uk/government/collections/growth-vouchers-programme>

² Clearly, this is a generalisation. One obvious point is that this tendency varies by policy area, with long-term outcomes being central to studies in education, for example.

others to revisit them and match interventions to longer-term outcomes. Initiatives like the ‘Datalab’ created in HM Revenue and Customs could allow this to happen.³ The Datalab works by listing the datasets available to researchers, inviting proposals for research and then allowing researchers to access and analyse the anonymised data. Thus, the data associated with any behavioural tax compliance initiative could be accessed later by other researchers.

It is reasonable to expect to see more studies of long-term effects emerge as the field matures. From BIT’s perspective, the need to focus more on ‘quick wins’ has subsided, and so we have more scope to pursue projects that pay off only in the longer term. In the meantime, we should also address the problem actively. For example, Frey and Rogers (2014) give a helpful overview of the different pathways that can be used to make long-term effects more likely. We can also identify at least three situations in which concerns over long-term effects may be avoided or assuaged.

One-off behaviours

Some interventions might have a ‘once and done’ property, and therefore have lasting effects without requiring any follow-up. Obvious examples here are those where an individual is being asked to complete a single action that changes their status in some way – for example, registering as a potential organ donor or enrolling in a workplace pension.

Resilient shifts

Another possibility is that an intervention may succeed in creating a new behavioural pattern that is sufficiently resilient to endure, even if the stimulus is withdrawn. Examples include when people are paid to visit the gym multiple times within a limited period, and then continue to show elevated attendance when payments are removed (Charness & Gneezy, 2009), or when a single letter leads to a sustained shift in the prescribing practices of doctors (Hallsworth *et al.*, 2016). The initial changes may be unintentional, rather than constituting a carefully constructed intervention. Data from the London Underground shows that during a two-day strike in 2014, most people found a new route to work. What is interesting is that 5% of these people did not return to their old commute: the disruption led to new sustainable practices (Larcom *et al.*, 2017).

³ <https://www.gov.uk/government/organisations/hm-revenue-customs/about/research#the-hmrc-datalab>

Environmental changes

An intervention might introduce a sustainable change to the decision-making environment that is likely to influence choices over the long term – for example, changes to the design and physical properties of a hospital waiting room. In this case, the behavioural stimulus endures, making it more likely that the behaviour will as well. A slight variation on this point are those interventions where a behavioural analysis suggests that the best option is *not to try to change behaviour*, but rather to mitigate its consequences. In other words, the intervention does not actually require people to do anything differently. Perhaps the best examples here are the successful interventions to reformulate food to reduce levels of salt and sugar while not changing consumer purchasing patterns.

Repeated exposure effects

A related concern is that repeated exposure to behaviourally informed approaches or interventions will lead to diminishing returns. We can distinguish between two main cases here: first, ‘structured repetition’, where the same approach is deliberately used as a direct follow-up to an initial intervention in order to reinforce its effects (e.g. reminders)⁴; and second, ‘unstructured repetition’, where individuals are exposed to the same kind of approach from different actors, at different times and in relation to different topics (e.g. when a social norm message appears in a different context). In both cases, the concern is that we may have a prior expectation that approaches become less effective with repeated exposure. The assumption underpinning this expectation is that the impact of an approach is driven by a novelty effect, wherein the approach: (a) succeeds in attracting attention in the first place; and (b) provides salient motivation to act in a particular way. Repeated exposure means the approach becomes less salient and novel, and thus less effective. (It is worth noting, therefore, that these concerns relate mainly to approaches that require attention from the individual, rather than those that are not immediately apparent – like default changes [see Hansen & Jespersen, 2013]. One would expect the latter to continue to be effective, since they rely more on automatic processes.)

Taking the ‘structured repetition’ question first, we can see that there is limited evidence available, which can lead us to compare wildly different studies. In some instances, we can see support for the diminishing returns hypothesis. For example, one BIT study showed that giving investment bankers sweets one year had a large impact on their willingness to donate a

⁴ Of course, this kind of follow-up may be unnecessary and not produce any marginal gains if the initial intervention succeeded in creating long-term effects.

day's salary to charity. The next year, the sweets had the same impact on bankers that had not received sweets before. However, bankers that had received the sweets the first time were less likely to give the second time around, although they remained more likely to give than participants who did not receive sweets at all (Sanders, 2015). At the same time, there is also evidence for an alternative hypothesis: repeated structured exposure can reinforce the initial intervention, rather than undermine it. Allcott and Rogers (2014) analysed the effect of repeated exposure to home energy reports containing social norm information and energy-saving tips. They found that such repeated exposure had the effect of keeping consumption lower compared to a group for whom the reports were dropped (although after two years, consumption remained lower in both groups than it was in the pre-treatment period). It seems clear that various aspects of the target behaviour and intervention could determine whether habituation or reinforcement occurs. We may hope that, as the number of studies that involve structured repetition increases, we are able to start drawing conclusions about what variables are associated with one outcome or another. We feel relatively confident that time will answer many of these questions.

On the other hand, we can be less sure about whether and how issues around unstructured repetition will be addressed, mainly because they are likely to be more complex and less tractable to analysis. Essentially, the problem of unstructured repetition is caused by success. BIT has a mission to increase the use of behavioural approaches that can achieve policy outcomes and increase the public good. As a result, we try to promote solutions that seem to work (as is common in academia). To give an example, Hallsworth *et al.* (2017) considered the impact of including social norm messages (“9 out of 10 people pay their tax on time”) on tax compliance and found a significant and positive impact. Similar effects of social norm messages have been replicated in other fields, some of which are reviewed in John *et al.* (2014).

The problem is that if the use of social norm messages spreads to many other policy domains, we have to start considering the effect of receiving the same or similar interventions from different quarters at the same time. How effective would social norm messages be if they were found on our gas bills, tax letters, inducements to travel by public transport and reminders to attend class at the local college? In other words, what are the general equilibrium effects? We might be concerned that the unstructured nature of the approaches, associating the same message with different behaviours, might mean that habituation occurs, but not reinforcement.

Unfortunately, most of these studies to date cannot answer this question, since they took place while few other studies were being conducted on the same sample at any given moment. Since the community of people running

such experiments has been small, it is likely that crossovers have been avoided. In the future, the main problem is likely to be one of coordination – since there are many different actors who could be using the same behavioural science approach, it is unlikely that any one study will be able to identify the other activities in play and take them into account. Perhaps the most we can do right now is to look at the studies that assess the impact of highlighting the use of behavioural science. Loewenstein *et al.* (2015) showed that emphasising to people that their behaviour had been influenced by a particular default setting did not undermine the effectiveness of the nudge (see also Bruns *et al.*, 2016). Clearly, this is not exactly the same issue, but it could suggest that a higher profile for behavioural policy may not necessarily mean a loss of effectiveness.

Problems with proxy measures

One response to the difficulties in including long-term effects is to use proxy measures.⁵ For example, many studies in education take attendance as an outcome measure (e.g. Kraft & Rogers, 2015; Chande *et al.*, 2017). We might reasonably expect that school attendance would have a positive impact on grades, as shown in Gottfried (2010)'s analysis of longitudinal data on school attendance and attainment. Similarly, one can infer long-term effects of behavioural interventions that reduce short-term unemployment (such as Altmann *et al.*, 2015) on the basis that they prevent 'employment scarring' – the tendency for one protracted spell of unemployment to lead to others, even if the first spell can be attributed to bad luck, such as graduating into a recession (Arulampalam *et al.*, 2001).

However, the causal relationship between the two types of measures may not be constant. This can be thought of as a behavioural science-specific form of the Lucas critique found in economics. Lucas (1976) argued that since consumption functions are not fixed and can respond to changes in circumstances, they are not 'policy invariant', and so may change in response to changes of policy. In economics, this critique remains important, as policies designed with a particular intended direction or magnitude of effect, based on current policy-response parameters in consumption functions, might be undermined because consumers' consumption functions change in response to the policy itself.

The equivalent argument is that the relationship between short-term and long-term behaviours is not 'nudge invariant': the act of nudging someone can cause their short- and long-term behaviours to become untethered from

⁵ Of course, proxy measures are needed for more than just estimating long-term outcomes. One may need them for outcomes that occur immediately but are difficult to measure.

each other. To illustrate, we can imagine that there is a well-established causal relationship between applying for jobs and leaving unemployment. It need not hold, however, that increasing job applications necessarily leads to increasing employment. For example, it is possible that marginal job applications are of a lower standard, or are less well targeted, than those that preceded them.⁶ Or we might speculate that a quick return to work may be less insulating from the psychological aspect of employment scarring if people are aware that they were nudged into work.

This problem is likely to be particularly acute with measures that are currently considered to be reliable proxies but where no causal chain exists. Self-control in children (e.g. measured by a ‘marshmallow test’) has been shown to predict several later-life outcomes and particularly later-life self-control (Mischel *et al.*, 1989). Given the existence of a problem early on in life, many would argue that it is at this stage that interventions should be targeted. Although this makes sense, we have no reason to expect (given the lack of a causal link between marshmallows now and alcohol later) that today’s proxy and tomorrow’s outcome are inextricably linked, or even linked that strongly. Hence, it may be possible that we can change one behaviour (self-control while young) without influencing the other(s) at all. This would not invalidate the intervention, as self-control when young may have all manner of direct positive benefits in the short and medium term. However, behavioural scientists should be cautious about extrapolating too far from their findings and should, through long-term data monitoring, attempt to identify whether relationships between variables that exist in the absence of interventions continue to do so in their presence.

Spillovers and unintended consequences

The argument about proxy measures can be taken in a more troubling direction. Rather than just questioning the relationship between immediate and proxy measures, we could point to the possibility that interventions are having unintended and unmeasured effects elsewhere. In this view, the focus in the experimental paradigm on specific, predefined outcome measures becomes a source of weakness as well as strength. Research on applying complex adaptive systems to policy emphasises how we may shift the dial on one outcome measure, even as problems mount elsewhere unobserved (Dörner, 1996; Dolphin & Nash, 2012).

⁶ We put to one side what have become standard arguments about steady states (Mortensen & Pissarides, 1999) and the lump-of-labour fallacy (Walker, 2007).

In some senses, this phenomenon may be unsurprising because governments may have a multitude of policy goals that they may be pursuing at the same time. Fisman and Golden (2017) offer a good example of the hidden trade-offs between these goals that interventions may reveal. They documented a US initiative that attempted to reduce fraud committed by grocery stores participating in a government food voucher scheme intended to improve nutrition. The fraud consisted of artificially inflating the prices charged for the goods covered by the scheme and was only possible because stores were not required to reveal the prices they charged to cash customers. The new initiative made changes that revealed the discrepancy. As hoped, the fraud disappeared. But the change also had the unintended consequence that many retailers simply stopped selling the high-nutrition foods altogether, since they were no longer profitable. Those who retained the foods increased their prices by nearly 10%. The result was poorer nutrition for both those on the programme and those in the local area who used the same stores.

Again, this argument about unintended consequences is not specific to behavioural policy interventions (see Cartwright & Hardie, 2012). But since behavioural science itself studies how such spillovers are triggered, there is a case to be made that it should be particularly sensitive to them. Here we can point to the large literature on how ‘licensing effects’ may mean that attempts at self-control in one domain create indulgence in another, how virtuous behaviour in one situation may increase dishonesty elsewhere and so on (Khan & Dhar, 2006; Merritt *et al.*, 2010).

True universals and cultural variation

As noted above, many more people and groups have been applying behavioural science to policy around the world. One description of what these people are trying to do is “to bring a more realistic model of human behaviour into ... policy and regulation” (Halpern, 2016). This is a useful shorthand and is not strictly speaking incorrect. But it also suggests that behavioural public policy has a tendency to treat human beings as a broad class, with variation occurring primarily at the individual level. Although we might, as in the case of stereotype threat, allow for the possibility of systematic differences by gender (Spencer *et al.*, 1999) and ethnicity (Steele & Aronson, 1995), behavioural scientists are relatively quiet on the issues of nationhood and culture.⁷

⁷ Again, this is a generalisation. We recognise that academics such as Croson and Buchan (1999) and Abeler *et al.* (2016) are exploring how trust games and honesty experiments produce different results in different countries.

As Henrich *et al.* (2010) noted, most people in the world are not ‘WEIRD’ – Western, educated, industrialised, rich and developed – but a large proportion of the people who apply behavioural science are (as are the subjects of their studies). This issue may be particularly acute for undergraduate participants in the labs of American universities, but it is still true to a great extent for participants on Amazon’s Mechanical Turk and for participants in many field experiments conducted by BIT, as these mostly take place in the UK, Australia and the USA.

This perspective makes it much more troubling that we lack field experimental evidence on the effectiveness of the same interventions in different contexts. We can make some general predictions, however. There is relatively strong evidence that Western societies have a stronger concept of individual, autonomous actors than East Asian societies, where a more collectivist perspective is often more prevalent (Heine, 2008). This difference may have an impact on the relative effectiveness of certain interventions. For instance, Bond and Smith’s (1996) study suggests that this difference means that social norm effects are larger in non-Western societies. Interestingly, Kettle *et al.* (2016, p. 36) found that social norms in tax letters were similarly effective in Guatemala, even though they stated that “65 percent of people pay their tax on time,” a very different number from that found on UK tax letter.⁸

The growth of teams around the world that draw on the same literature and develop similar interventions may allow us to address this challenge. If new teams are able to conduct very similar replication studies, it could be possible to establish which (if any) heuristics and biases are common to which societies and, for those that are not universal, what factors moderate or mediate their effects. This process is likely to identify issues, contexts and cultures that challenge behavioural scientists to develop and test new theories and interventions.

‘Reverse impact’

In recent years, the UK higher-education sector has been encouraged to document how its research has ‘impact’, defined as “an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia” (Penfield *et al.*, 2014, p. 21). Interestingly, behavioural scientists working on policy issues may experience the situation in reverse. Rather than publishing peer-reviewed research that may then influence government action, they may alter government actions

⁸ An obvious point to make is that the crucial factor here may be whether the social norm updates the individual’s beliefs in a positive or negative way. In other words, although 65% may seem low, it may be significantly higher than the recipient expected.

and then attempt to publish the results in peer-reviewed journals. In other words, the impact comes first.

The issue is that there may be several barriers to academic publication, many of which BIT has experienced in its history and tried to surmount. Publication is usually entirely dependent on the approval of senior officials, who may see little reason to approve this request, particularly if there is any transfer of data involved (even anonymised data). The issues are particularly acute if the people responsible for the study are public officials, as BIT staff members were until spring 2014. In this case, there are often few resources provided to support the publication process, which is likely to be seen as a luxury. But even if resources are supplied by academic partners, there may be differences over the proposed timing, framing and conclusions of any potential publication. While BIT has always been pressing the case for publication, it has not always been successful.

Should we be concerned by the lack of ‘reverse impact’? Here, we should distinguish between initiatives that are not made public and those that are made public but do not go through the additional step of peer-reviewed publication. The former create obvious problems in terms of reducing the transparency of government, as well as potentially creating a ‘public file drawer’ problem if it is null results that are selectively held back. The latter are less serious but may mean that (depending on exactly how they have been reported) the quality and reliability of their results cannot be fully assessed. The worry is that if not all the studies have been prepared and structured carefully, this may introduce quality uncertainty, where even good studies are suspected of being potentially flawed (Akerlof, 1970). One obvious solution for this second problem is for public but non-peer-reviewed reports to provide enough detail to allow a reasonable judgement of their quality.

The replication crisis

Finally, we wanted to mention the challenge that the replication crisis brings behavioural public policy. BIT takes many of its intervention ideas from existing lab studies or field experiments and is therefore dependent on their reliability. If this reliability is questionable, there are two main consequences for behavioural public policy: first, and most importantly, it would mean that we have wasted resources trying to implement concepts that are not viable – resources that could have been allocated more profitably; and second, it could damage the trust policy-makers, and the public, have in behavioural science, with the result that they become reluctant to use the approach in future.

This is a challenge that will not go away immediately, since new questionable practices are still being discovered regularly. International efforts on pre-registration, data sharing, identification of statistical irregularities and replication consortia (such as the Many Labs Replication Project; see Klein *et al.* 2014) are all part of the emerging solution to the crisis. However, as behavioural scientists, we are all responsible for helping the field overcome this challenge. In particular, those behavioural scientists who are helping to form policy (and thus determine the allocation of limited public resources) have a special responsibility to review the relevant evidence critically and in detail, bearing in mind the problems that have recently emerged. Robustly evaluating the ensuing interventions becomes even more important. Finally, and most painfully, we need to accept with humility that some of our own trials may have been one-off findings. The implication is that we should be attempting to replicate an intervention as a matter of course before trying to scale up an intervention. This would present a significant challenge, since the public sector is often still reluctant to evaluate interventions in the first place.

Cluster 2: Opportunities

Behavioural government

Those who apply behavioural science to policy have focused most of their energies on using new approaches to improve policy outcomes. Much less attention has been paid to behavioural science as a tool to improve the way government itself functions (Lodge & Wegrich, 2016). This division of resources reflects how incentives are structured: individuals and groups using behavioural science may not seek to challenge fellow public sector actors because they need them as allies for implementing interventions. Yet there are strong reasons for doing so.

The main reason is that these critiques are almost certainly correct. Government *is* vulnerable to biases and pathologies in the way it acts; no serious observer could argue otherwise. Moreover, many of these traits are the same ones that behavioural science has discussed extensively in relation to individuals. A minister visiting a hospital or factory may be exposed to a particular piece of information that acts as a breeding ground for confirmation bias. A finance officer will rely heavily on the previous year's budget as the default when allocating resources, even if 'zero-based budgeting' is applied.⁹

⁹ 'Zero-based budgeting' is a process in which all expenditure is built up from a 'base' of zero, rather than taking last year's budget as a starting point. The idea is that this removes any anchoring effects and allows more careful consideration of any potential expenditure.

There are real gains to be made here; the task should not be dismissed as too difficult or mundane. Behavioural scientists have much to add to existing studies on organisational behaviour in this regard.

The other reason is more tactical. Critics of using behavioural public policy often use the fact that government is also vulnerable to biases to justify the claim that governments should simply refrain from nudging (or other applications) altogether (Waldron, 2014). This does not necessarily follow: a stronger argument is that this points towards the need for *more* behavioural science, not less, just applied to government itself. Yet, in general terms, behavioural scientists have allowed this argument to go mostly unchallenged, partly because they have not developed the arguments or generated the examples to do so. With both these reasons in mind, BIT is currently working on a self-funded project on this issue.

Scaling interventions

Scaling interventions represents a specific opportunity for taking a behavioural approach to government itself. The application of behavioural science has often adopted a basic approach of ‘experiment and then scale’: a trial is conducted on a sample (which nonetheless might be relatively large) and the best-performing variant is adopted more widely. In many cases, interventions are constructed to ensure that they can be integrated into existing large-scale practices, so they have a clear path to wider adoption.

Perhaps the best examples are those that concern changes to messages that are already being sent out, since the change itself can require little expenditure. Hence, many of BIT’s studies have concerned modification of letters (e.g. Hallsworth *et al.*, 2017), text messages (e.g. Haynes *et al.*, 2013) or web forms (Kettle *et al.*, 2017). Other kinds of intervention are a feasible route to scaling up, even if they are not as cheap or easy as message changes. Examples might be changes to the way payments are structured, such as using lotteries (SBST, 2016), different match rates (Huck & Rasul, 2011) or social incentives (BIT, 2016).¹⁰

However, despite these apparent advantages, the proportion of behavioural science interventions that reach scale could be much higher. Even when interventions are scaled, this does not necessarily mean that they retain their original effectiveness. A well-known example concerns surgical checklists, a patient safety measure that requires the surgical team to check whether certain

¹⁰ Of course, other high-profile policies informed by behavioural science have needed substantial investment to achieve scale. In the UK, opt-out defaults for pensions, as studied by Choi *et al.* (2004), required the creation of the National Employment Savings Trust (NEST).

actions have been carried out. Checklists had been highly successful in improving patient outcomes in previous studies, reducing surgical complications from 11% to 7% (Haynes *et al.*, 2009). However, when their use was mandated across hospitals in Canada, that effect disappeared, despite self-reported compliance being over 90% (Urbach *et al.*, 2014). Checklists also did not improve outcomes in a hospital in The Netherlands; patients for whom the checklists were used correctly had improved outcomes, but they were only used for 39% of the procedures (Van Klei *et al.*, 2012). In these cases, both context and how the checklist was implemented, as well as a potential lack of intensive training or monitoring of compliance, may have affected their effectiveness (Urbach *et al.*, 2014).

It is important to note that the failure of innovations to reach scale is often observed in the public administration literature: the issue is not specific to behavioural science interventions (Bason, 2010). Nevertheless, there are at least two reasons why failure to scale is particularly relevant to behavioural public policy. The first is that recent efforts in this field have tended to couple behavioural science with an emphasis on the importance of experimentation (Haynes *et al.*, 2012; SBST, 2015; Ames & Hiscox, 2016). There are many good reasons for this approach. However, the danger is that it means too much focus is placed on trials and trial results as ends in themselves (particularly in the light of our comments about proxy measures above). Furthermore, it is arguable that academic incentives encourage innovative results and highly cited publications, but fail to encourage researchers to scale their interventions. We need to ensure that the individuals and organisations applying behavioural science to policy consider scalability when selecting projects and that they are incentivised to look beyond the successful completion of a trial. Luckily, this is something government and other organisations (e.g. the Educational Endowment Fund and the Health Foundation) are becoming increasingly interested in, providing new opportunities for trials in this area.

The second reason is that one could argue that the issue of scaling is, at heart, one of organisational behaviours – and thus behavioural science itself may have something useful to say about the problem. We are aware, of course, of the contribution that implementation science has made in this regard. But we argue that more of the attention and resources dedicated to behavioural public policy should be directed to this question (as for government activity in general). For example, the literature on diffusion of innovations has established that homophily (i.e. the tendency for individuals to associate with similar others) plays a significant role in how ideas do or do not reach scale (e.g. Yavaş & Yücel, 2014). It seems likely that behavioural science could contribute to this field by, for example, offering insights into how the source of a

message affects its persuasiveness or into how social norms function. But we are still some way from using behavioural science to give policy-makers recommendations about how to use homophily to maximise the likelihood that others take up their innovation. How best to scale up and spread interventions is an area where there is a strong need for further research that uses the same robust methods that are associated with behavioural science trials.

Social diffusion

Many applications of behavioural science to policy have adopted a simple, uni-directional model of influence: a public sector actor attempts to influence (usually) an individual, organisation or group. This approach is clearly important and covers many common policy situations. However, it notably excludes avenues for influence that exist *between* individuals, organisations or groups. Sacerdote's (2001) canonical study of social influence, in which undergraduates at Dartmouth College were randomly assigned to roommates, found that spatial proximity (and, presumably, the relationships that form as a consequence) significantly influences Grade Point Average (GPA) at the end of students' first year of college. Moreover, it is likely that this estimate is a lower bound, since most friendships are formed through associative matching rather than random assignment.

Arguably, behavioural science's focus on individual decision-making means it has neglected relevant insights from theories about social networks and systems thinking (Ormerod, 2010). Obviously, this is a generalisation, and many such valuable studies do exist (e.g. Drago *et al.*, 2015; Kim *et al.*, 2015). But the potential gains to policy-makers are so large that they justify more work. In a world of limited resources, a better understanding of how to harness peer-to-peer transmission of behaviour could mean that the same or better outcomes are achieved at much less cost (since far fewer contacts from government would be needed). The same is true if governments are attempting to limit this transmission.

BIT has conducted some preliminary studies on how to create 'network nudges'. One example compared the difference between: asking an email recipient (at an investment bank) to click on a link to donate to charity; asking the recipient to "reach out and email their friends and colleagues"; and asking them to reach out and tell their acquaintances "about the huge contribution their donation can make." For each of these variants, the proportion of investment bankers making a donation was 12.4%, 23.6% and 38.8%, respectively (BIT, 2016). This is clearly a very simple example, however, and much more work is needed to identify the best ways to use social networks.

Nudging organisations

Most behavioural science interventions have focused on individuals, most commonly in their roles as citizens and consumers, and less frequently, but increasingly, as employees. There has been much less work that targets organisations, and there are a few likely reasons why. One is that (generally, but not exclusively) the main academic disciplines informing behavioural policy have tended to use the individual as their unit of analysis. Another is that regulators, which constitute one of the main points of contact between government and organisations, have traditionally taken a legalistic approach that is averse to experimentation and innovation. Recent developments show that this attitude is changing (Financial Conduct Authority, 2013). In addition, the drive for robust evaluation may have guided attention towards individuals. Since there are fewer organisations than people, randomised trials that use organisations as the unit of analysis may be more likely to encounter problems with statistical power.

The more important point is that many policy-relevant decisions are made by organisations. The majority of fossil fuels are consumed by organisations, through transportation of products, the powering of factories and so on. Although consumer behaviour is an important part of preventing and mitigating climate change, any strategy that does not incorporate changing business behaviour is destined to fail. This argument also applies to the products that firms make. Policy interventions may try to reduce carbon emissions by getting people to drive less (Leape, 2006) or to reduce calorie consumption using food labels (Roberto *et al.*, 2010). However, in both cases, the problem could be solved upstream if firms made cars that used less petrol or reformulated their products to contain fewer calories – something the UK's new sugared drinks levy will aim to address.

How far can we translate findings from individual psychology to organisational psychology, or is this attempt hopelessly naïve? After all, businesses are not people, but they *are* made up of people.¹¹ If the government were to write a letter to all grocery stores asking them to stock more low-calorie drinks and fewer high-calorie ones, that letter, if it is read at all, will be read by a person, and so to the extent that people can be nudged, we could expect businesses to be 'nudgeable'.

This argument may not follow, however. Many firms may have created processes to prevent single employees from making large decisions – indeed, these processes may have been created explicitly to guard against cognitive biases.

¹¹ Again, we are aware that the field of organisational behaviour has produced many important and valuable contributions to these questions, not least Simon (1947).

Two or more decision-makers may not be more biased than one, but if only one of them has been nudged, the intervention's effect may be reduced. There is also a question of identity (Akerlof & Kranton, 2000). When people are working for a firm, we do not expect them to behave the way that they do at home. The evidence suggests that relatively weak, short-term prompts that invoke a particular identity, such as being 'a voter' rather than simply 'voting' (Bryan *et al.*, 2011), can have substantive effects on subsequent behaviour. Employees may have even more intense and sustained exposure to prompts that encourage behaviours that are in line with corporate identities (Akerlof & Kranton, 2005). Finally, we must acknowledge the great variety of organisational forms. For example, it is plausible that smaller firms may have less of a distinct identity from the individual(s) running them, which could lead them to behave more like individuals. Since organisation size is often recorded, there is a strong case for trying to answer this question by gathering together existing evidence on how size interacts with behavioural science interventions.

Governments have often approached the behaviour of organisations through the lens of regulation. At this stage, we do not know enough about how behavioural science can improve, complement or even replace regulation, despite the many valuable contributions of Cass Sunstein (2011). BIT has been considering this question, from the 2010 MINDSPACE report – which stresses that its framework “does not attempt to replace [legislation and regulation] ... it extends and enhances them” (Dolan *et al.*, 2010, p. 49) – to its recent report on applying behavioural insights to regulated markets (Costa *et al.*, 2016). However, it is clear that we have made much less progress here than we have in the domain of individual decisions. Oliver (2015)'s discussion of 'budging' rightly emphasises the potential gains that can be made in the field of behavioural regulation; clearly, more work is needed here.

Thorny problems

One criticism of behavioural policy is that the breadth of its usefulness is limited to certain domains where one-off behaviours with binary desirable decisions are prevalent, such as tax compliance, attending appointments or enrolling in a pension plan. It is certainly the case that the majority of government teams that work in behavioural science have typically focused their early work on changing these kinds of behaviours. We (Hallsworth & Sanders, 2016; Sanders & Halpern, 2016) have been among those advocating for pursuing so-called 'low-hanging fruit' early in the life of a behavioural insights team. The social value of this work has now, become easy to downplay,

particularly when policy-makers have become accustomed to ‘tax letter trials’ producing effects.

However, we argue it would be disappointing if compliance were the only application of behavioural science active in policy ten years from now. Our hope is that behavioural science has as much to offer on some trickier, more complex problems. When we think of a particularly challenging area of government policy, we often think of areas where the traditional tools of regulation, incentives and information have been extensively utilised and yet the problem persists. At first, this can seem something daunting and to be avoided, but it is precisely here that new tools should be used. If we consider that behavioural economics came into existence to explain phenomena that standard economic analysis found inexplicable, perhaps behavioural science can solve problems that standard economic tools have found insoluble.

The last few years have seen the beginnings of these changes at BIT in the UK. We are making dedicated attempts to consider how behavioural science can be applied to relieve poverty (Gandy *et al.*, 2016); to curb recidivism, where financial and regulatory incentives are ineffective for a substantial number of people; and to prevent corruption and bribery, which may be so culturally pervasive as to evade traditional tools. Work in these areas is still in its early stages. As a community of practice, we must not become disheartened if/when we do not achieve immediate successes in these areas. Where we do start to make progress, we must guard against the temptation to overstate our successes and to declare prematurely that the battle has been won.

Summary and conclusion

Tackling these challenges will require sustained effort and new collaborations. We should be prepared for the fact that interventions over the next few years will be harder to implement and may produce more ambiguous effects. As the replication crisis rolls onwards, we must similarly be ready for the possibility that many more things that we had previously taken to be true will cease to be credible.

We should, however, remain optimistic. A lot of progress has been made in a relatively short space of time. Breaking economists’ intellectual hegemony over government was never going to be straightforward, but – as is already apparent – it is worth the effort.

Acknowledgements

We are grateful to colleagues at the Behavioural Insights Team for numerous conversations over the years that have informed the content of this paper.

References

- Abeler, J., D. Nosenzo and C. Raymond (2016), 'Preferences for truth-telling', *The Centre for Decision Research and Experimental Economics, University of Nottingham*.
- Akerlof, G. (1970), 'The market for lemons: Qualitative uncertainty and the market mechanism', *Quarterly Journal of Economics*, **84**(3).
- Akerlof, G. A. and R. E. Kranton (2000), 'Economics and identity', *The Quarterly Journal of Economics*, **115**(3): 715–753.
- Akerlof, G. A. and R. E. Kranton (2005), 'Identity and the economics of organizations', *The Journal of Economic Perspectives*, **19**(1): 9–32.
- Allcott, H. and Rogers, T. (2014), 'The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation', *American Economic Review*, **104**(10): 3003–37.
- Altmann, S., F. Armin, S. Jäger and F. Zimmermann (2015), 'Learning about job search: A field experiment with job seekers in Germany', *IZA Institute of Labor Economics*.
- Ames, P. and M. Hiscox (2016), Guide to developing behavioural interventions for randomised controlled trials: Nine guiding questions.
- Arulampalam, W., P. Gregg and M. Gregory (2001), 'Unemployment scarring', *The Economic Journal*, **111**(475): 577–584.
- Bason, C. (2010), *Leading public sector innovation: Co-creating for a better society*, Policy Press.
- Behavioural Insights Team (2016), Update report: 2015–16. Available online <http://www.behaviouralinsights.co.uk/uncategorized/the-behavioural-insights-teams-update-report-2015-16/>
- Bond, R. and P. B. Smith (1996), 'Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task', *Psychological bulletin*, **119**(1): 111.
- Bruns, H., E. Kantorowicz-Reznichenko, K. Klément, M. Luistro Jonsson and B. Rahali (2016), Can nudges be transparent and yet effective?
- Bryan, C. J., G. M. Walton, T. Rogers and C. S. Dweck (2011), 'Motivating voter turnout by invoking the self', *Proceedings of the National Academy of Sciences*, **108**(31): 12653–12656.
- Cartwright, N. and J. Hardie (2012), *Evidence-based policy: A practical guide to doing it better*, Oxford University Press.
- Chande, R., M. Luca, M. Sanders, X. Soon, O. Borcan, N. Barak-Corren, E. Linos, E. Kirkman and S. Robinson (2017), Increasing attendance and attainment among adult students in the UK: Evidence from a field experiment.
- Charness, G. and U. Gneezy (2009), 'Incentives to exercise', *Econometrica*, **77**(3): 909–931.
- Choi, J. J., D. Laibson, B. C. Madrian and A. Metrick (2004), 'For better or for worse: Default effects and 401 (k) savings behavior', *University of Chicago Press: Perspectives on the Economics of Aging*, 81–126.
- Costa, E., K. King, R. Dutta, and F. Algate (2016), Applying behavioural insights to regulated markets.
- Croson, R. and N. Buchan (1999), 'Gender and culture: International experimental evidence from trust games', *The American Economic Review*, **89**(2): 386–391.
- Della Vigna, S. (2009), 'Psychology and economics: Evidence from the field', *Journal of Economic literature*, **47**(2): 315–72.
- Dolan, P., M. Hallsworth, D. Halpern, D. King and I. Vlaev (2010), MINDSPACE: influencing behaviour for public policy.
- Dolphin, T. and Nash, D. (Eds.), (2012), *Complex New World: translating new economic thinking into public policy*, Institute for Public Policy Research.
- Dörner, D. (1996), *Logic of Failure: Why Things Go Wrong and What We Can Do to Make Them Right*, New York: Metropolitan Books.

- Drago, F., F. Mengel, and Traxler, C. (2015), Compliance behavior in networks: Evidence from a field experiment.
- Financial Conduct Authority (2013), Applying behavioural economics at the Financial Conduct Authority. *Occasional Paper No.1*.
- Fisman, R. and M. Golden (2017), 'How to fight corruption', *Science*, 356(6340): 803–804.
- Frey, E. and T. Rogers (2014), 'Persistence: How treatment effects persist after interventions stop', *Policy Insights from the Behavioral and Brain Sciences*, 1(1): 172–179.
- Gandy, K., K. King, P. Streeter Hurle, C. Bustin and K. Glazebrook (2016), Poverty and decision-making: How behavioural science can improve opportunity in the UK.
- Gottfried, M. A. (2010), 'Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: An instrumental variables approach', *American Educational Research Journal*, 47(2): 434–465.
- Hallsworth, M., J. List, R. Metcalfe and I. Vlaev (2017), 'The behaviorist as tax collector: Using natural field experiments to enhance tax compliance', *Journal of Public Economics*, 148: 14–31.
- 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.
- Hallsworth, M., T. Chadborn, A. Sallis, M. Sanders, D. Berry, F. Greaves and S. C. Davies (2016), 'Provision of social norm feedback to high prescribers of antibiotics in general practice: A pragmatic national randomised controlled trial', *The Lancet*, 387(10029): 1743–1752.
- Halpern, D. (2015), *Inside the Nudge Unit: How small changes can make a big difference*, Random House.
- Halpern, D. (2016), It's time to bring more realistic models of human behaviour into economic policy and regulation, *Bank Underground*.
- Hansen, P. G. and A. M. Jespersen (2013), 'Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy', *European Journal of Risk Regulation*, 4(1): 3–28.
- Haynes, A. B., T. G. Weiser, W. R. Berry, S. R. Lipsitz, A. H. S. Breizat, E. P. Dellinger, ... and A. F. Merry (2009), 'A surgical safety checklist to reduce morbidity and mortality in a global population', *New England Journal of Medicine*, 360(5): 491–499.
- Haynes, L. C., D. P. Green, R. Gallagher, P. John and D. J. Torgerson (2013), 'Collection of delinquent fines: An adaptive randomized trial to assess the effectiveness of alternative text messages', *Journal of Policy Analysis and Management*, 32(4): 718–730.
- Haynes, L., B. Goldacre and D. Torgerson (2012), Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials. *Cabinet Office*.
- Heine, S. J. (2008), *Cultural Psychology*, New York: W. W. Norton.
- Henrich, J., S. J. Heine and A. Norenzayan (2010), 'Most people are not WEIRD', *Nature*, 466(7302): 29.
- Huck, S. and I. Rasul (2011), 'Matched fundraising: Evidence from a natural field experiment', *Journal of Public Economics*, 95(5): 351–362.
- Institute for Government and Cabinet Office (2010), MINDSPACE: Influencing behaviour through public policy.
- John, P., M. Sanders and J. Wang (2014), The use of descriptive norms in public administration: A panacea for improving citizen behaviours?
- Kettle, S., M. Hernandez, M. Sanders, O. Hauser and S. Ruda (2017), Failure to CAPTCHA attention: Null results from an honesty priming experiment in Guatemala.
- Kettle, S., M. Hernandez, S. Ruda and M.A. Sanders (2016), 'Behavioral interventions in tax compliance: Evidence from Guatemala'.

- Khan, U. and R. Dhar (2006), 'Licensing Effect In Consumer Choice', *Journal of Marketing Research*, 43(2): 259–266.
- Kim, D. A., A. R. Hwang, D. Stafford, D. A. Hughes, A. J. O'Malley, J. H. Fowler and N. A. Christakis (2015), 'Social network targeting to maximise population behaviour change: a cluster randomised controlled trial', *The Lancet*, 386(9989): 145–153.
- Klein, R. A., K. A. Ratliff, M. Vianello, R. B. Adams, Jr, Š. Bahník, M. J. Bernstein, ... and Z. Cermalcar (2014), 'Investigating variation in replicability: A "many labs" replication project', *Social psychology*, 45(3): 142.
- Kraft, M. A. and T. Rogers (2015), 'The underutilized potential of teacher-to-parent communication: Evidence from a field experiment', *Economics of Education Review*, 47: 49–63.
- Larcom, S., F. Rauch and T. Willem (2017), 'The benefits of forced experimentation: Striking evidence from the London underground network', *The Quarterly Journal of Economics*, 132(4): 2019–2055.
- Leape, J. (2006), 'The London congestion charge', *The Journal of Economic Perspectives*, 20(4): 157–176.
- Lodge, M. and Wegrich, K. (2016), 'The rationality paradox of nudge: Rational tools of government in a world of bounded rationality', *Law & Policy*, 38(3): 250–267.
- Loewenstein, G., C. Bryce, D. Hagmann and S. Rajpal (2015), 'Warning: You are about to be nudged', *Behavioral Science & Policy*, 1(1): 35–42.
- Lucas, R. E., Jr. (1976), 'Econometric policy evaluation: A Critique', In Brunner, K. and A. H. Meltzer (eds.), *The Phillips Curve and Labor Markets, Carnegie-Rochester Conference Series on Public Policy. Volume 1, Journal of Monetary Economics, Supplement*, 19–46.
- Merritt, A.C., D. A. Effron, and B. Monin (2010), 'Moral self-licensing: when being good frees us to be bad', *Social and Personality Psychology Compass*, 4(5): 344–357.
- Mischel, W., Y. Shoda and M. L. Rodriguez (1989), 'Delay of gratification in children', *Science*, 244(4907): 933.
- Mortensen, D. T. and C. A. Pissarides (1999), 'New developments in models of search in the labor market', *Handbook of Labor Economics*, 3: 2567–2627.
- OECD (2017), Behavioural Insights and Public Policy: Lessons from Around the World. Available online <http://www.oecd.org/gov/regulatory-policy/behavioural-insights-and-public-policy-9789264270480-en.htm>
- Oliver, A. (2013), 'From nudging to budging: Using behavioural economics to inform public sector policy', *Journal of Social Policy*, 42(4): 685–700.
- Oliver, A. (2015), 'Nudging, shoving, and budging: Behavioural economic-informed policy', *Public Administration*, 93(3): 700–714.
- Ormerod, P. (2010), *N Squared: Public policy and the power of networks. RSA Pamphlet.*
- Penfield, T., M. J. Baker, R. Scoble and M. C. Wykes (2014), 'Assessment, evaluations, and definitions of research impact: A review', *Research Evaluation*, 23(1): 21–32.
- Pykett, J., R. Jones and M. Whitehead (eds.) (2017), 'Psychological governance and public policy: Governing the mind, brain and behaviour', *Oxford: Routledge.*
- Roberto, C. A., P. D. Larsen, H. Agnew, J. Baik and K.D. Brownell (2010), 'Evaluating the impact of menu labeling on food choices and intake', *American Journal of Public Health*, 100(2): 312–318.
- Sacerdote, B. (2001), 'Peer effects with random assignment: Results for dartmouth roommates', *The Quarterly Journal of Economics*, 116(2): 681–704.
- Sanders, M. (2015), In search of the limits of applying reciprocity in the field: Evidence from two large field experiments.
- Sanders, M. and D. Halpern (2016), 'Nudging by government: Progress, impact, and lessons learned', *Behavioral Science and Policy*, 2(2): 52–65.

- Sanders, M. and E. Kirkman (2014), 'I've booked you a place. Good luck: A field experiment applying behavioural science to improve attendance at high-impact recruitment events', *The Centre for Market and Public Organisation Working Paper Series*, 14: 334.
- Shafir, E. (ed.), (2013), *The behavioral foundations of public policy*, Princeton University Press.
- Simon, H. (1947), *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*. Macmillan.
- Social and Behavioural Science Team (SBST) (2015), Annual report. Available online <https://sbst.gov/download/2015%20SBST%20Annual%20Report.pdf>
- Social and Behavioural Science Team (SBST) (2016), Annual report. Available online <https://sbst.gov/download/2016%20SBST%20Annual%20Report.pdf>
- Spencer, S. J., C. M. Steele and D. M. Quinn (1999), 'Stereotype threat and women's math performance', *Journal of experimental social psychology*, 35(1): 4–28.
- Steele, C. M and J. Aronson (1995), 'Stereotype threat and the intellectual test performance of African Americans', *Journal of Personality and Social Psychology*, 69(5): 797.
- Sunstein, C. R. (2011), 'Empirically informed regulation', *The University of Chicago Law Review*, 78 (4): 1349–1429.
- Sunstein, C. R. (2014), *Simpler: The Future of Government*, Simon and Schuster.
- Thaler, R. (2015), *The Making of Behavioral Economics: Misbehaving*, New York.
- Thaler, R. H. and C. R. Sunstein (2008), *Nudge: Improving decisions about health, wealth and happiness*, New Haven: Yale University Press.
- Urbach, D. R., A. Govindarajan, R. Saskin, A. S. Wilton and N. N. Baxter (2014), 'Introduction of surgical safety checklists in Ontario, Canada', *New England Journal of Medicine*, 370(11): 1029–1038.
- Van Klei, W. A., R. G. Hoff, E. E. H. L. Van Aarnhem, R. K. J. Simmermacher, L. P. E. Regli, T. H. Kappen, ... and L. M. Peelen (2012), 'Effects of the introduction of the WHO "Surgical Safety Checklist" on in-hospital mortality: a cohort study', *Annals of Surgery*, 255(1): 44–49.
- Waldron, J. (2014), *It's All for Your Own Good*, New York Review of Books.
- Walker, T. (2007), 'Why economists dislike a lump of labor', *Review of Social Economy*, 65(3): 279–291.
- Whitehead, M., R. Jones, R. Howell, R. Lilley and J. Pykett (2014), *Nudging All Over the World: Assessing the Global Impact of the Behavioural Sciences on Public Policy. Economic Social & Research Council Report*. Swindon, UK.
- World Bank Group (2015), *Mind, Society, and Behavior*. World Development Report. Washington, DC: The World Bank.
- Yavaş, M. and G. Yücel (2014), 'Impact of homophily on diffusion dynamics over social networks', *Social Science Computer Review*, 32(3): 354–372.