

Nudging by government: Progress, impact and lessons learnt

David Halpern & Michael Sanders

abstract

“Nudge units” within governments, most notably in the United Kingdom and the United States, seek to encourage people to behave a certain way by using insights gained from behavioral science. The aim is to influence people’s choices through policies that offer the right incentive or hurdle so that people choose the more economically beneficial options. Getting people to save for retirement, eat more healthful foods, and pay their taxes on time are some examples of institutionally desirable activities. The 10-fold rise in “nudge” projects undertaken since 2010—more than 20 countries have deployed or expressed interest in them—have revealed many lessons for policymakers. Chief among these lessons: the necessity of obtaining buy-in from key political leaders and other stakeholders, and the benefits of testing multiple intervention strategies at once. Although detailed cost–benefit analyses are not yet available, we estimate that behaviorally inspired interventions can help government agencies save hundreds of millions of dollars per year.

Core Findings

What is the issue?

The U.K. government's Behavioral Insights Team (BIT) is delivering monetary benefits in the region of hundreds of millions of dollars, if not billions. To replicate this success, other governments must first work to successfully embed behavioral interventions in the policy mix.

How can you act?

Selected best practices include:

- 1) Building strong relationships with academia through cross-disciplinary advisory panels
- 2) Starting with rapid, low-cost, multi-arm behavioral trials using existing administrative data
- 3) Investing in impact valuations to measure the return on investment from interventions

Who should take the lead?

Policymakers in government, academics working in behavioral science

“There is not enough money for retirement” is a common lament among workers and policymakers alike. As things stand now, the U.S. Social Security trust fund will run empty by 2035,¹ and about half of all Americans have saved less than \$10,000 for their golden years.² In the past decade, policymakers have tackled this failure of people to act now for a better tomorrow by redirecting people’s own natural inertia. Specifically, more and more organizations require employees to opt out of retirement plans rather than opt in, as in the past. In the United Kingdom alone, the opt-out approach has meant more than 5 million extra workers have started saving for their workplace pensions since 2012. By the end of 2016, the default rule change reached the entire population of United Kingdom workers, including small firms and even micro-employees (people who work only a few hours for a given employer, often a family) such as nannies and cleaners.

The success of increasing retirement savings shows the value of behavioral interventions. Since the publication of the book *Nudge: Improving Decisions About Health, Wealth, and Happiness* by Richard Thaler and Cass Sunstein in 2008³ and especially over the last few years, governments have increasingly incorporated overtly behavioral approaches into policy. Of course, almost all government policy is a form of behavioral influence, insofar as it aims to influence the actions of human beings through either legislation, regulation, or the provision of information. However, policymakers have moved toward getting people to change their behavior. It is an overt acceptance, or even embracing, of behavioral science in the form of behavioral economics, psychology, and related fields, as a tool for adjusting people’s behavior.

In this article, we review developments in and the expanded use of behavioral science by governments and other institutions. We also tentatively estimate the number of government-conducted randomized controlled trials (RCTs) that explicitly attempt to apply findings from these fields and offer the beginnings of a profile of their impacts. Finally, we reflect on early lessons learned, particularly for the benefit of policymakers and academics in the process of building

this capability into their own governments. Our analysis is not a comprehensive overview but instead draws directly on our own experiences and knowledge, particularly of the U.K. government’s Behavioural Insights Team (BIT), which serves as a model that many other governments have begun to follow.

A Brief History

Governments have long drawn on tacit knowledge of human behavior to shape how their citizens act. However, in the early 2000s, governments on both sides of the Atlantic began to more overtly incorporate psychological and behavioral factors into policy, regulation, and program delivery. Thaler and Sunstein’s article⁴ on libertarian paternalism attracted the attention of U.S. policymakers, while in the United Kingdom, the idea of applying behavioral science came to the attention of government officials after *Personal Responsibility and Changing Behaviour*⁵ was published from within the Prime Minister’s Strategy Unit, which existed during Tony Blair’s administration to provide advice and policy analyses.

In the United States, the 2008 publication of *Nudge*³ and the subsequent move of one of its authors, Cass Sunstein, into an influential position within the White House in 2009 gave a major boost to embedding behavioral approaches into policy. As head of the Office of Information and Regulatory Affairs and with the support of President Obama, Sunstein was able to intervene on a range of regulatory issues, particularly through the use of executive orders. For instance, these orders enabled the Environmental Protection Agency to regulate greenhouse gas emissions and set fuel economy standards without congressional approval. (Sunstein left the Obama administration in 2012.)

The United Kingdom soon followed the U.S. example. Although the 2004 Prime Minister’s Strategy Unit paper sparked negative political and media reactions,⁶ the newly elected coalition government in 2010, partly inspired by the perceived impact of *Nudge* on Barack Obama’s presidential campaign and administration, created 10 Downing Street’s BIT.

An important but subtle difference emerged between the British and American approaches at this time. Whereas Sunstein primarily relied on the use of executive orders to incorporate behavioral approaches into policies, the U.K. unit pursued a more experimental approach, one that resembled Lockheed Martin's Skunk Works programs, where engineers are not assigned to specific projects with a short-term goal but instead are given greater freedom to pursue innovative and novel ideas, the expectation being that even if most of these ideas fail, the successes will more than pay for the unit's costs.

The advantage of Sunstein's approach was that it offered the prospect of large and immediate effects by instantly or quickly transforming entire domains. The disadvantage, of course, is that executive orders often lack the legacy of congressional approval. Therefore, the orders may have only short lives and face dissolution by court challenges (such as rulings issued by the U.S. Supreme Court against President Obama's orders on immigration) or by new executive orders from a different administration (such as President Trump's first-day order to begin dismantling the Affordable Care Act). The United Kingdom's more modest approach often involves running small-scale trials to test interventions inspired by behavioral science. This more experimental approach brings with it other advantages, not least being that it builds up an evidence base that can ultimately prove highly persuasive to an otherwise skeptical audience of senior public servants and commentators. This approach has since been replicated overseas, notably in the White House Social and Behavioral Sciences Team (SBST).

Detailed accounts have recently documented the struggle to get the U.S. academic and policy communities to engage with behavioral science⁷ and the history of the United Kingdom's BIT.⁶ For now, we simply note that two linked strands of activity have emerged. First, policymakers using behavioral approaches have sought to incorporate a more realistic account of human behavior in their work, for example, in the way consumer energy markets must provide information about their tariffs. As recently highlighted by Stanford

University economist Raj Chetty, this strategy can lead to new policy proposals, better predictions, and a different perspective on the relative efficacy of existing policy tools.⁸ Second, behavioral approaches have brought in their wake, at least in their U.K. manifestation, a form of "hyper-empiricism," in that variations in interventions are constantly being tested and their causal impacts are continually estimated. Halpern has termed the approach *radical incrementalism*: although each intervention on its own may seem modest, when the approach is applied widely and persistently, it is transformative. The road to this stage has not been entirely smooth, and much has been learned en route. In the next section, we articulate a few of the keys to successfully applying behavioral science to policy and some of the lessons learned on the way.

Key to Successful Nudging: APPLES

Attempts over the last decade to bring behavioral science out of the laboratory and into the world of policy have produced many lessons. Policymakers seeking to create "nudge units" within their own government or other public bodies are advised to pay heed to the following necessary components that can be summarized in the simple mnemonic of APPLES: administrative support, political support, people, location, experimentation, and scholarship.⁶ We outline APPLES in greater detail below.

Administrative Support

Ensure you have senior level buy-in inside the system. For BIT, it was key that we had the support of the cabinet secretary, the United Kingdom's most senior government official, and that he personally agreed to chair BIT's steering board. His backing and participation sent a powerful signal to the rest of government and gave us leverage when we needed it, especially because other permanent secretaries (a *permanent secretary* being a department's most senior-ranking civil servant) were more skeptical. For these doubters, showing them the early results of BIT's tax letter trials that upped tax payments by an estimated £20 million⁹ was the first step in winning their support. (See Table 1 for more details on the tax letter intervention.)

Table 1: Examples of the United Kingdom's behaviorally based interventions & their reach

Intervention	Reach	Impact
Change to opt-out saving for workplace pensions (from 2012, starting with larger employers)	27 million employees	An increase of 9 million people newly saving or saving more in qualifying workplace pensions by 2018 is expected as a result of automatic enrollment. ^A 5.4 million extra savers enrolled by August 2015, before extension to smaller firms.
Tax prompts to encourage timely payment , such as adding the line "most people pay their tax on time" in letters to taxpayers	10.4 million eligible for self-assessment, and particularly those who are late to file or pay	An estimate from 2012 of early trials of Her Majesty's Revenue & Customs and the Behavioural Insights Team (BIT) was that \$300 million was brought forward. The estimate has not been formally updated, although scale and reach have subsequently expanded substantially. ^B
Job search improvements to get people back to work faster , by revising processes and prompts targeting those out of work and on benefits (for example, advisers use an implementation intention intervention to prompt job seekers to set out what, when, and how they will be looking for work in the coming week)	Codified and rolled out to 25,000 Jobcentre advisers in 2014, reaching around 800,000 people at any one time on Jobseeker's Allowance or Universal Credit (the United Kingdom's working-age social security program) ^C	Days on benefits have been reduced by an estimated 5 million to 10 million, ^D based on effect sizes found in a regional stepped-wedge trial. This equates to state welfare cost savings of \$75 million to \$150 million per annum (excluding wider benefits to job seekers).
E-cigarette availability : BIT advice starting in 2011 led to the decision to ensure widespread availability of electronic cigarettes in the United Kingdom (although sales to those under 18 years of age were banned)	More than 9 million smokers in the United Kingdom ^E	2.8 million smokers, or ex-smokers, now use e-cigarettes in the United Kingdom. In 2015, Public Health England estimated e-cigarettes to be 60% more effective as a route to quitting than rival methods, and e-cigarettes have become the most dominant route to quitting smoking in the United Kingdom. ^F
Organ donation : prompts to encourage people to join the organ donor register, added at the end of car-tax payment bills, based on the result of an eight-arm BIT trial	20 million people a year	Some 96,000 extra donors joined the register per annum.
Reduction in unnecessary antibiotic prescriptions , through letters to 20% of the highest-prescribing general medical practitioners	12 million people, covered by 13,000 of the highest-prescribing general medical practitioners	A 3.3% reduction in antibiotic prescriptions in the target population was sustained at 6 months, equating to just under a 1% reduction in prescriptions nationally. Excess antibiotic prescriptions likely propel the rise of resistant bacterial strains, seen as the greatest medical threat to the current generation by the United Kingdom's chief medical officer.

A. National Audit Office. (2015). Automatic Enrolment to workplace pensions, <https://www.nao.org.uk/wp-content/uploads/2015/11/Automatic-enrolment-to-workplace-pensions.pdf>

B. Behavioural Insights Team. (2016). Update report 2015–16. Retrieved from <http://www.behaviouralinsights.co.uk/publications/the-behavioural-insights-teams-update-report-2015-16/>

C. Office for National Statistics. (2016). *UK labour market: February 2016* [Statistical bulletin]. Retrieved from <http://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/uklabourmarket/february2016>

D. Halpern, D. (2015). *Inside the nudge unit: How small changes can make a big difference*. London, United Kingdom: WH Allen.

E. Action on Smoking and Health. (2016). Fact sheets. Retrieved from <http://ash.org.uk/information/facts-and-stats/fact-sheets>

F. Public Health England. (2015). *E-cigarettes: A firm foundation for evidence-based policy and practice*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/454517/E-cigarettes_a_firm_foundation_for_evidence_based_policy_and_practice.pdf

Political Support

Consider how the approach fits with the political narrative and instincts of the governments concerned. Interest in behavioral approaches from Prime Minister David Cameron and Deputy Prime Minister Nick Clegg, as well as their close aides, facilitated BIT's launch in 2010.

People

Create a team with the right mix of skills and expertise. At least as important as subject experts are people with the battle-hardened experience of working in government and large organizations. Personal relationships with those whom you will need as allies are equally important. As a result, there is no single type of person that BIT has recruited. BIT could not function well if it did not have team members who each possessed at least one of six key skill sets: understanding of government, knowledge of behavioral science, knowledge of policy and intervention design, analytical skills, interpersonal communication skills, and management skills.

We endeavor to create project teams comprising individuals who, through their academic training and professional experience, have a combination of these skills. For example, many BIT employees came from the U.K. Civil Service or had careers in other governments or international organizations such as the United Nations; these individuals have extensive knowledge of how governments work and how policy is designed. Other employees have come straight from academia and typically are at the doctoral or postdoctoral level. They provide the analytical expertise as well as knowledge of behavioral science and intervention design. Our colleagues who are former practitioners, such as National Health Service managers, teachers, and coaches from Jobcentre Plus (the main U.K. government service that maintains direct contact with unemployed job seekers and administers out-of-work benefits), provide further understanding of intervention design and public service delivery. Finally, we have many employees who have come from management consulting and other professional-services firms and whose strengths lie in management, delivery of public services, and communication. The combination of skills

"Embrace empirical methods."

makes individual project teams greater than the sum of their parts.

Location

Choose a location close to the institutions and people with whom you wish to work rather than a fancy office 20 minutes away. So much of government, as of life, is about being in the right place at the right time. In certain places, people regularly bump into each other and conduct impromptu business. In the United Kingdom, such well-trafficked places include the lobby of 10 Downing Street, Parliament, and even on the street of Whitehall, a U.K. government thoroughfare.

Experimentation

Embrace empirical methods. You'll need to demonstrate to skeptics and fence sitters that your new approach works, and you will need to quantify its impact. But, more fundamentally, you should follow the logical progression of test, learn, adapt. Behavioral science is especially well suited to experimental approaches, as they often involve minor changes to existing processes rather than the initiation of new processes.

Scholarship

Know the behavioral literature and details of the challenges you will face. Most everybody has some everyday knowledge of psychology, but you need a team that contains people with detailed, expert-level knowledge of the field—either through professional experience and practice or advanced study—who are plugged in to the latest thinking and results. Identify your local and relevant academic experts and form an advisory group.

Seven Specific Lessons

We also learned more specific lessons. Although APPLES provides a high-level framework, the following seven lessons offer more practical, day-to-day advice, gleaned from our experiences of developing, implementing, and scaling behaviorally informed trials within the U.K. government.

1. Use Rapid, Low-Cost Trials That Apply Existing Administrative Data Gathered by the Government.

BIT's work on tax collection and payments of fines^{9,10} provides good examples of this approach. Over the course of several years, BIT conducted a series of trials,^{9,11,12} both large and small, in which late-paying taxpayers were sent notices from the tax authority. They received modified versions of the standard letters, each applying a different insight from behavioral science. Tax collectors routinely and automatically send out millions of letters each year, so modifying the content of the letters is inexpensive, and the outcome—whether people pay their tax and how much—is something that the administration already records. Trials are much easier to conduct and attract more administrative support when they are targeted on an outcome or objective that a government department already has. This also makes it more likely that a positive result will be adopted and scaled up. It can be a sensible trade-off to make a few methodological and measurement compromises on such exploratory trials if the compromises make the trials less burdensome to administrators, with promising results then leading to larger, more robust trials. For example, a small pilot study was run to evaluate a new behaviorally informed process in Jobcentre Plus. Although imperfect, the pilot offered strong enough evidence of effectiveness to convince key decisionmakers to invest in a larger, more robust trial.¹³

2. Get Field-Worker Input. User-centered design principles and qualitative research are important tools that can be used to develop hypotheses and iron out problems in prototype interventions before a full trial is begun. When working with Jobcentre Plus to redesign the experience of new out-of-work benefit claimants, BIT conducted extensive qualitative fieldwork prior to designing the intervention. In doing so, BIT staff experienced firsthand both the claimant's journey and the challenges faced by the Jobcentre Plus coaches who help people get back to work. Through this experience, they found that initial Jobcentre Plus meetings looked backward, focusing on what the job seeker had done in the last two weeks rather than on what the job seeker was going to do. Moreover, job seekers had to sign 14 documents on their first visit, leaving almost no time

to discuss employment. The intervention that was ultimately tested¹³ drew heavily on that fieldwork, which revealed processes that could not have been properly observed or understood from behind a desk in Whitehall.

An example of the counterfactual can be found in an unpublished study that Michael Sanders conducted in 2012. The goal was to replicate a 2004 study by Thaler and Benartzi,¹⁴ which found that employees much prefer a gradual increase in their charitable donation rate over an ultimately smaller, one-time increase in their donation. The intervention design suffered from what we call *theory-induced blindness*, where testing a specific theory is the sole concern of the researcher, who becomes unable to see anything else. In the end, the intervention was a failure, significantly reducing the number of donations made. In hindsight, this result was obvious. The intervention was delivered via e-mail, a medium ill-suited for conveying an idea as complicated as precommitting to escalating giving rates. Because the intervention design considered only the theory and not the end users or the context, it was fatally flawed.

3. Prepare Yourself for Failure. Often an idea that looks good in theory or seems like it will be effective when conceived within a central government office does not succeed as expected when it is implemented in the messiness of the real world. This may be a result of optimism among policymakers themselves.¹⁵ As Sunstein noted in his 2014 book *Why Nudge? The Politics of Libertarian Paternalism*, for every bias identified for individuals, there is an accompanying bias in the public sphere.¹⁶ Recognizing this fact and attempting to identify and overcome our own biases and preconceptions through extensive fieldwork and challenging ideas within the BIT has, in our view, helped us to generate better interventions. Just as important, practitioners who work in a field every day will rarely hesitate to tell you when they think that an intervention will fail; their opinions should be taken seriously.

4. Consider the Ease of Scalability in Intervention Design. Interventions that are simple and inexpensive to implement, even if they have small absolute effects, may be more cost-effective than

impressive but complex interventions. Adding one line to a tax letter that raises payment rates by 5 percentage points, or 15%, may seem modest compared with redesigning the tax system, but the cost-effectiveness is very great. Similarly, BIT's work on voter registration¹⁷ found that offering a £5,000 lottery incentive to register to vote was only slightly more effective than a £1,000 lottery incentive. Offering a smaller prize was therefore more cost-effective and wrought less political controversy, and the monetary savings allowed agencies to deploy the strategy more widely. Similarly, small-scale interventions, such as the use of implementation intention booklets for job seekers, are much easier and cheaper to scale than more intensive programs that involve more active employment support, even if the absolute effect per person may be modest.

5. Be Pragmatic & Err on the Side of Multifaceted or Bundled Interventions in Early Trials. If a positive impact is found, subsequent RCTs can disentangle the effects of individual elements of the intervention. For example, when working on improving attendance in community colleges, BIT tested a suite of text messages, which were delivered at regular intervals.¹⁸ The results were impressive: the text-message interventions reduced dropout numbers at the end of the first semester by one third. But the design of the experiment meant that it was not possible to isolate the active ingredient in the success.

6. Choose Multi-Arm Trials Over Single-Arm Trials. Comparing more than one intervention with a control tends to be preferable on both methodological and political grounds to single-arm trials. From a policy perspective, behavioral scientists should be concerned with not just whether a given intervention changes an outcome in the desired way but whether it works better than other possible interventions. In our 2016 study,¹⁹ we found that charitable donations by staff at an investment bank increased if the bank requested that their manager ask the staff to donate. In the multi-arm trial, we could test the best way for the manager to make the request. The best prompt increased the proportion donating from 5% to 35%; the worst prompt only increased the rate to 8%. Without testing multiple arms simultaneously, we would not have

made this discovery. Politically, it is also much easier to make the case for a multi-arm trial. It's much easier to sell the definitive message: "Minister, we'll find out which version works best at producing a desired result" (multi-arm trial), as opposed to chin-down message of, "Minister, we'll be able to conclusively show if the program flopped" (single-arm trial).

7. Walk Before You Run, Even if That Means Leaving Your Passion Project for a Later Date.

It is often better to start with modest interventions (or combinations of interventions) or, at least, those that have been rigorously tested elsewhere to establish your expertise and a baseline of trust with the administration. Your dream intervention will probably involve a lot more than sending a text message or a letter, and you'll likely struggle to get a complex trial off the ground without establishing an initial trust bank with policymakers. We quite often take a long list of interventions to policymakers at the first meeting. Many times, our top-priority interventions are tossed out almost immediately, because they either are too complicated to implement, deviate too much from established practice, or are deemed "too wacky."

Together, these seven lessons embody a pragmatic approach to using behavioral science trials in government policymaking, and these recommendations arguably stand in some contrast to the more purist approach of conducting randomized controlled trials in the academic world. A dose of pragmatism may necessitate statistical corrections for imperfections in design and underscore the need to use convergent evidence to interpret results. However, they also tend to lead to a more reliable path to policy impact and allow for the testing of academic theory in a much tougher and more demanding real-world context.

The Impact of Behavioral Approaches

Although interest in applying behavioral approaches to policy has increased over the last decade, honesty dictates that we admit that interest doesn't necessarily translate into impact: many ideas are fashionable for a time, then



close to half of all Americans currently have **\$10,000** saved for retirement

\$75-150m

estimated direct savings in cash benefits from a BIT intervention in job-seeking activity

2%

reduction in energy use per individual consumer once consumers were told how efficient they were relative to each other

vanish without a trace. Most governments have, at best, a rather patchy record of measuring the impact of their policies in a systematic and reliable way.²⁰ Furthermore, when impacts are measured, the causal ancestry—how the policy was developed, whose idea it was, and who gave the necessary authorizations and funding—of any given policy can be hard to reconstruct. There's truth to the saying that success has many parents, but failure is an orphan.

Here we offer four relatively simple, linked ways that may demonstrate the power of behavioral approaches to help build effective policies, limit political quagmires, and benchmark global adoption of such approaches.

- Widespread impacts in the United Kingdom and United States
- Spread across countries
- Volume of behavioral study trials and policy interventions
- Estimate of intervention impact, monetized in US dollars

Widespread Impacts in the United Kingdom and United States

Behavioral scientists can certainly claim their interventions have touched the lives of tens of millions of people (see Table 1). Changes in pension saving rules are perhaps the most obvious behaviorally inspired intervention on both sides of the Atlantic in the last decade. We've outlined the impact of implementing savings defaults in the United Kingdom and in the United States. Work by John Beshears, currently at the Harvard Business School, has demonstrated the huge impact on savings enrollment rates that even firm-level defaults can have.²¹

Interest Among Other Countries

By 2013, a number of other governments and public bodies had started to become interested in applying behavioral science to policy (see the sidebar *Spread of Behavioral Science Programs around the Globe*).

Two early movers were Australia (New South Wales in particular) and Singapore. Both set up behavioral teams in central governments. The combination of the ideas articulated in books such as *Nudge*³ and the steady stream of practical trial results from the United Kingdom's BIT was particularly intriguing to pragmatic public administrators in these countries.

By 2015, interest and active application had spread to many other countries, typically by government treasuries and tax administrations that saw how BIT's small, low-cost interventions could boost tax collection totals. In 2014, the German government announced that it was setting up a small team inside the Chancellery, with direct links to Chancellor Angela Merkel. In Italy, the government of Prime Minister Matteo

Spread of Behavioral Science Programs Around the Globe

2013

Australia (New South Wales), Singapore: Behavioural Insights Unit established the Department of Premier and Cabinet with the secondment of Rory Gallagher from the U.K. Behavioural Insights Team to the Department of Premier and Cabinet. Singaporean Public Services Division and Ministry of Manpower begin randomized trials.

United States: White House launches the Social and Behavioral Sciences Team, similar to the United Kingdom's Behavioural Insights Team. Similar teams exist in the governments of some cities such as New York and Chicago.

United Kingdom: Civil service reform document calls for the adoption of behavioral science strategies by all government departments.

2014

Germany: Announced it would set up a small team inside the Chancellery, with direct access to Chancellor Angela Merkel.

Italy: Prime Minister Matteo Renzi's office published a document on modern policymaking that outlines the relevance of nudging.

European Commission: Announced creation of a behavioral unit inside the European Commission's Joint Research Centre.

Netherlands: Network of departmental teams is established.

2015

Australia (Federal Government): Team headed by Harvard Professor Michael Hiscox established in the Department of the Prime Minister and Cabinet.

2016

Australia (Victoria): Team established in the Department of Premier and Cabinet.

references

Renzi published a document on modern policymaking that outlined the relevance of nudging.²² And within the European Commission, long seen as a bastion of traditional regulation, it was announced that a behavioral unit would be created within the Commission's Joint Research Unit.

In 2013, the White House launched its own Social and Behavioral Sciences Team (SBST), headed by Maya Shankar, a young neuroscientist out of Stanford. This new team brought into the Obama administration many of the same methods that had characterized the U.K.'s BIT. Although SBST's genesis and activities are independent of those of the BIT, the similarity of methodology can be clearly seen, for example, in their first report.²³

The attendance roster at the September 2015 Behavioural Exchange conference, hosted in London, demonstrates the range of governments using or considering behavioral approaches to policy. The 900 delegates and speakers included officials and advisers from more than 20 countries. Nations beyond those listed above that are actively considering behavioral approaches include Canada, the Netherlands, Sweden, Denmark, Israel, Ireland, Mexico, the United Arab Emirates, Jamaica, and Brazil.

International bodies are also tapping the behavioral and experimental approaches to policy. They include the European Commission, the World Bank, the United Nations Development Programme, the Organisation for Economic Co-operation and Development, and the World Economic Forum. These institutions are actively supporting the spread of the use of behaviorally-inspired approaches into Central and Latin America, Central and Eastern Europe, Africa, and South Asia through direct investments, trials, summits, and publications such as the World Bank's World Development Report titled *Mind, Society, and Behavior*.²⁴

Within countries, the spread is being facilitated by the movement of people among influential roles. For example, it is no coincidence that the Australian state of Victoria created its own behavioral

insights capability after the appointment of Chris Eccles as secretary of the Department of Premier and Cabinet—a role he had held in New South Wales, where he had seen the results from its behavioral insights team. Meanwhile, both the United Kingdom and the United States are seeing significant uptake of behaviorally based trials by local, regional, and city governments.

Building Strong Bridges Between Government and Academia

The growth in number and scope of government behavioral insights teams has been supported by a strong sense of collaboration, both among teams in different countries and with academia. The collaborations have taken various forms. BIT maintains an academic advisory panel to provide guidance on the team's work, and established BIT's research fellow program that recruits promising doctoral students to work with BIT for short stints (some fellows continue with BIT, while others return to academia with experience of government). The collaborations on special projects with academics outside the BIT are bearing published fruit.^{9,10,17,18,25–27}

These special projects are born from two main formats. In one, a long-standing relationship between an academic and BIT leads to specific projects that naturally align with each other's interests. The second, more common format involves discussions between BIT members and academics to establish areas of interest and who's researching what, so when something appropriate comes up, those working on behavioral issues in government will know whom to contact.

Volume of Trials

An arguably more solid measure of adoption of behavioral approaches is the number of trials being initiated by governments. No simple database documents this. BIT is seeking to follow emerging good practice of publishing outline protocols of trials and the results of these trials on a regular basis,^{28,29} but even in the United Kingdom, practical and political pressures sometimes prevent trials from being made available in the public domain.

“Behavioral scientists can certainly claim their interventions have touched the lives of tens of millions.”

Table 2. Estimated number of trials conducted by behavioral units in government (2010–2016)

Country	Number of trials	Primary source institution
United Kingdom	300–400	Behavioural Insights Team (Cabinet Office), Her Majesty's Revenue & Customs (tax collection), Education Endowment Foundation, Financial Conduct Authority
United States	30–50	Social and Behavioral Sciences Team (White House Office for Science and Technology Policy)
Australia	10–25	Department of Premier and Cabinet (New South Wales), VicHealth, Department of Premier and Cabinet (Victoria)
Singapore	20–30	Ministry of Manpower, Prime Minister's Office
The Netherlands	5–15	Treasury; Department for Infrastructure and the Environment; Ministry of Business Affairs
International	10–30	World Bank, United Nations Development Programme
Total	375–550	

Using a combination of public material and our own knowledge, Table 2 gives an indication of the number of behaviorally based trials under way across governments and public bodies.

We estimate that around 375–550 behaviorally inspired trials have been explicitly and intentionally initiated by governments over the last 5 years. The majority of these trials were initiated in the last 2 years, indicating an acceleration in activity. This is likely an underestimate, considering the strategy of using behavioral insights while crafting policies has recently spread to state, municipal, and city governments, which makes counting harder. All of this research promises to lead to an explosion of new results.

These numbers do not include the wider rise in the use of trial methods in general. For example, the United Kingdom's Educational Endowment Foundation, set up by the Department for Education in 2011 to systematically test and identify what works in education, has sponsored and supported around 100 large-scale trials involving more than 4,000 schools and over 600,000 children. Only a minority of these trials have an overt

link to the behavioral literature, whereas others test more conventional interventions, such as whether student grades can be increased by hiring teaching assistants or paying students for performance. Such interventions are excluded from our estimates.

Impact Valuations: The Return on Investment

Table 1 shows some examples from the United Kingdom of the reach of a selection of behaviorally based interventions. For some of them, an estimate of impact is relatively straightforward. For example, we can derive an estimate of effect size from a regional stepped-wedge rollout of a BIT intervention to encourage job seekers to plan their job search activity. The original single-site intervention suggested that job seekers exposed to the implementation intention intervention were around 10% more likely to be off benefits, presumably because they had returned to work (although this was not confirmed), after 13 weeks, but this single-site intervention had several problems, such as displacement effects (for example, one group of job seekers might be simply getting jobs faster but taking them from others in the same area). Another possibility is that the heavy involvement of the BIT team might have led to originator effects that would not be seen in a wider rollout. The stepped-wedge multisite design, in contrast, required the codification of the intervention and training-by-trainer implementation, and it had much less possibility of displacement of effects. Unsurprisingly, the regional trial led to an effect size that was considerably smaller, at around 1.7 percentage points, but that provides a reasonable estimate of the likely effect size when the same codification and standardized training were expanded to the national level. In this case, this leads to an estimate of direct savings in cash benefits of around \$75–\$150 million, not factoring in the wider economic benefits of a more active labor market and reduced emotional scarring of individuals who reenter the workforce faster.

Of course, the sample size in a trial, the reach of the intervention, and even the effect size are not by themselves an indication of impact in real policy terms. For example, one of BIT's

trials on organ donation involved a sample of over 1 million people to test eight variations of prompts asking people if they would join the (voluntary) organ-donation register. Adding the prompt to the annual car tax renewal process resulted in around a quarter of a million new donors joining the register each year. Adopting the most effective of the eight variants contributed approximately 100,000 extra registrations to this total. These are large numbers, but it is important to recognize that even adding an extra 100,000 donors is likely to save only a few lives a year, valuable though these are.

Similarly, there's no doubt that changing pension defaults in the United Kingdom has led to massive increases in savings—certainly running to billions of pounds since 2012. Yet, it's difficult to calculate the scale of the economic benefit that follows. For example, some have argued that it might have been better over this particular economic period to have stimulated extra consumer spending rather than saving. The most obvious benefit of increased savings ought to be that governments would use the success of automatic enrollment to wind back tax subsidies to consumers or firms. A 2014 study published in the *Quarterly Journal of Economics*³⁰ estimated that the net effect of a \$100 tax subsidy encouraging people to save is a mere \$1 of extra pension saving by consumers. To date, however, governments have been wary of winding down these subsidies—the cost of which is estimated to run to more than \$30 billion for the United Kingdom alone and much more than that in the United States—for fear of destabilizing the pension market or of political backlash among high-turnout voters.

Sustainability

One key challenge is to take the important results from current trials to scale. For example, the United Kingdom's tax office, Her Majesty's Revenue and Customs, has taken a true test-and-learn approach, where the results of small-scale trials—that is, small in terms of sample size and complexity—have been expanded to become national policy even as new, novel tests continue. In general, successful small-scale RCTs need to be extended to the broader population if the

“One key challenge is to take the important results from current trials to scale.”

true potential of behavioral science in policy is to be realized.

Another challenge is to look at the longer-term effects of interventions. This challenge comes in two parts: habits and habituation. Habits concern the ability of behavioral interventions to have lasting effects on people's lives by making them change their behavior not just immediately after an intervention but in the longer term. As Frey and Rogers³¹ pointed out, the evidence that currently exists is fairly limited and often not promising. Although short-term effects may be sufficient to identify a bias or other phenomena in an academic setting, in a policy context, more work is clearly needed. The second component of these long-term effects concerns habituation, or what happens when people are exposed repeatedly to the same kind of behavioral intervention. This is an area that warrants significant study as these interventions become more commonplace.

Finally, it is worth noting that some of the most effective interventions may come from the private sector. The strategy of giving consumers feedback about how much energy they use relative to their more efficient neighbors (declarative social norms)—notably promoted by Opower, an energy services firm—has been rolled out to more than 50 million consumers so far, and that number is rising. Although the 2% reduction in energy use per individual consumer this intervention averages³² may appear modest, when aggregated across all 50 million Opower customers, this is a big impact. Also note that in some cases, an intervention may be more appropriately run by an entity other than the government: when governments are not the best actors to intervene, charities or corporations can sometimes get an intervention to the target population at scale.

We estimate (conservatively, we think), on the basis of more precise examples such as the Jobcentre Plus trial and scale-up or tax trials, that the monetary benefits of behavioral interventions are safely estimated in the hundreds of millions of dollars. If more wide-ranging estimates are used, such as including a monetized value for years of life saved by a particular strategy (for example, if someone is persuaded to use e-cigarettes rather than smoking actual cigarettes) or estimating the combined benefits of taxes paid through repeated trials, then the benefits almost certainly run to many billions of dollars.

Frustratingly, at least an order of magnitude difference exists between these conservative and wider-ranging estimates. Of course, this disparity is not unique to behaviorally based interventions. The historic rarity of RCTs and robust evaluations in most policies, as well as the complexity of estimating effects—for example, did military or security spending actually prevent an attack?—mean that only a tiny proportion of the trillions spent by governments across the world can be said to have been subject to a meaningful cost-benefit impact analysis. The promotion of RCTs and related research methods by behavioral scientists may start to change this landscape through, for example, the rise of What Works Centres and a growing understanding that experimental methods can give pragmatic and rapid answers to operational and policy questions that policymakers and the public want answered.

Measuring Success: A Final Summary

Over the last decade, behavioral approaches have moved from being an interesting idea to increasingly mainstream practice within the policy community. The quest for impact is still very much a work in progress, both to identify tomorrow's equivalent of the default changes on pensions and to scale up the promising interventions that are currently being studied.

This importance of buy-in from stakeholders cannot be overstated. It goes to the heart of the APPLES mnemonic, which emphasizes that there is no single component for success.

Enthusiasm from politicians can only translate into policy triumphs if the machinery of government can also be convinced that your idea is going to work. Teams of just academics or just policymakers are less likely to be successful, as both ingredients are necessary to successfully conceive, test, and implement policies that influence behavior. Perhaps most important, the bar for evaluation has been raised over the last few years, as organizations like the Education Endowment Foundation have led the way in showing what "good" looks like, and this is set to continue. Applying lessons from science to policy without rigorous testing is not desirable—nor is it easy to get away with. With the right combination of skills and infrastructure, the future is bright for policymakers or academics looking to apply behavioral science to policy.

author affiliation

Halpern, Behavioural Insights Team; Sanders, Behavioural Insights Team and Blavatnik School of Government, Oxford University. Corresponding author e-mail: Michael.sanders@bi.team

author note

We are grateful for the insights from friends and colleagues in the Behavioural Insights Team, in particular to Owain Service and Michael Halls-worth, and to Craig Fox, Sim Sitkin, and three anonymous reviewers for thoughts on a draft of this article. Thanks are also due myriad collaborators on randomized trials described briefly herein who are too numerous to mention and to Chiara Varazzani, Aisling Ni Chonaire, and Ariella Kristal for research assistance.

references

1. Social Security and Medicare Board of Trustees. (2016). A summary of the 2016 annual reports. Retrieved March 5, 2017, from <https://www.ssa.gov/oact/trsum/>
2. Kirkham, E. (2016). 1 in 3 Americans have \$0 saved for retirement. Retrieved from GOBankingRates website: <https://www.gobankingrates.com/retirement/1-3-americans-0-saved-retirement/>
3. Thaler, R. H. & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.
4. Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *American Economic Review*, 93(2), 175–179.
5. Halpern, D., Bates, C., Mulghan, G., & Aldridge, S. (with Heathfield, A., & Beales, G.). (2004). *Personal responsibility and changing behaviour: The state of knowledge and its implications for public policy* [Issue paper]. London, United Kingdom: Cabinet Office, Prime Minister's Strategy Unit.
6. Halpern, D. (2015). *Inside the nudge unit: How small changes can make a big difference*. London, United Kingdom: WH Allen.
7. Thaler, R. H. (2015). *Misbehaving: The making of behavioral economics*. New York, NY: Norton.
8. Chetty, R. (2015). Behavioral economics and public policy: A pragmatic perspective. *American Economic Review*, 105(5), 1–33.
9. Hallsworth, M., List, J., Metcalfe, R., & Vlaev, I. (2014). *The behavioralist as tax collector: Using natural field experiments to enhance tax compliance* (NBER Working Paper No. 20007). Cambridge, MA: National Bureau of Economic Research.
10. Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2015). *The making of homo honoratus: From omission to commission* (NBER Working Paper No. 21210). Cambridge, MA: National Bureau of Economic Research.
11. Cabinet Office Behavioural Insights Team. (2012). *Applying behavioural insights to reduce fraud, error and debt*. Retrieved from http://38r8om2xjhhl25mw24492dir.wpengine.netdna-cdn.com/wp-content/uploads/2015/07/BIT_FraudErrorDebt_accessible.pdf
12. Haynes, L. C., Green, D. P., Gallagher, R., John, P., & Torgerson, D. J. (2013). Collection of delinquent fines: An adaptive randomized trial to assess the effectiveness of alternative text messages. *Journal of Policy Analysis and Management*, 32, 718–730.
13. Gallagher, R., Gyani, A., Halpern, D., Hanes, S., Kirkman, E., Reinhard, J., . . . Service, O. (2016). *Applying behavioural insights to the labour market: Evidence from two randomised trials* [Working paper]. London, United Kingdom: Behavioural Insights Team.
14. Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow: Using behavioral economics to increase employee savings. *Journal of Political Economy*, 112(Suppl. S1), S164–S187.
15. Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39, 806–820.
16. Sunstein, C. R. (2014). *Why nudge? The politics of libertarian paternalism*. New Haven, CT: Yale University Press.
17. John, P., MacDonald, E., & Sanders, M. (2015). Targeting voter registration with incentives: A randomized controlled trial of a lottery in a London borough. *Electoral Studies*, 40, 170–175.
18. Chande, R., Luca, M., Sanders, M., Soon, X.-Z., Borcan, O., Barak-Corren, N., . . . Robinson, S. (2015). *Curbing adult student attrition: Evidence from a field experiment* (Harvard Business School Working Paper 15-065). Cambridge, MA: Harvard University.
19. Sanders, M., & Norton, M. (forthcoming). *Network nudges: Leveraging social networks for charitable donations in two field experiments*
20. Morse, A. (2010). *Taking the measure of government performance*. Retrieved from National Audit Office website: <https://www.nao.org.uk/wp-content/uploads/2010/07/1011284.pdf>
21. Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2011). Behavioral economics perspectives on public sector pension plans. *Journal of Pension Economics and Finance*, 10, 315–336.
22. Renzi, M., & Giannini, S. (2014). *La buona scuola: Facciamo crescere il paese* [The good school: We grow the country]. Retrieved from Passo dopo passo website: <http://passodopopasso.it/wp-content/uploads/lbs-web.pdf>
23. Social and Behavioral Sciences Team. (2015). *Annual report*. Retrieved from <https://sbst.gov/download/2015%20SBST%20Annual%20Report.pdf>
24. World Bank. (2015). *World development report 2015: Mind, society and behavior*. Washington, DC: Author.
25. Sanders, M. (2017). Social influences on charitable giving in the workplace. *Journal of Behavioral and Experimental Economics*, 66, 129–136. doi:10.1016/j.socsc.2015.12.004
26. Sanders, M., & Smith, S. (2016). Can simple prompts increase bequest giving? Field evidence from a legal call centre. *Journal of Economic Behavior & Organization*, 125, 179–191.
27. John, P., Sanders, M., & Wang, J. (2014). *The use of descriptive norms in public administration: A panacea for improving citizen behaviours?* Retrieved from SSRN website: https://papers.ssrn.com/sol3/papers2.cfm?abstract_id=2514536
28. Behavioural Insights Team. (2015). *Update report 2013–15*. Retrieved from http://www.behaviouralinsights.co.uk/wp-content/uploads/2015/07/BIT_Update-Report-Final-2013-2015.pdf
29. Behavioural Insights Team. (2016). *Update report 2015–16*. Retrieved from <http://www.behaviouralinsights.co.uk/publications/the-behavioural-insights-teams-update-report-2015-16/>
30. Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H., & Olsen, T. (2014). Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark. *Quarterly Journal of Economics*, 129, 1141–1219.
31. Frey, E., & Rogers, T. (2014). Persistence: How treatment effects persist after interventions stop. *Policy Insights From the Behavioral and Brain Sciences*, 1(1), 172–179. doi:10.1177/2372732214550405
32. Fox, C. R., & Sitkin, S. B. (2015). Bridging the divide between behavioral science & policy. *Behavioral Science & Policy*, 1(1), 1–12.