A MANIFESTO

for Applying Behavioral Science

THE BEHAVIOURAL INSIGHTS TEAM

Michael Hallsworth

ACKNOWLEDGMENTS

Thanks to Lila Tublin for drafting and editing support, Richard O'Brien for communications support, Dilhan Perera for thoughts on the "Predict and Adjust" section, and Alex Gyani for his input on the "Replication, Variation, Adaptation" section.

Thanks to S. Banerjee, E. Berkman, A. Buttenheim, F. Callaway, J. Collins, E. Costa, J. Doctor, D. Halpern, P. John, S. Kousta, T. Marteau, M. Muthukrishna, C. Payne, D. Perrott, K. Ruggeri, R. Schmidt, D. Soman, H. Strassheim, C. Sunstein, and members of the Behavioural Insights Team for their feedback on previous drafts. I am grateful to the four peer reviewers at Nature Human Behaviour for their helpful suggestions.

SUMMARY

This is a manifesto for how applied behavioral science can fulfill its true potential.

The behavioral insights movement has flourished over the last decade.[1] There is now a vibrant ecosystem of practitioners, teams, and academics building on each other's work across the globe. Their focus on robust evaluation means we know that this work has had an impact on important issues such as antimicrobial resistance, educational attainment, climate change, and obesity.

The Behavioural Insights Team is proud to have been a pioneer of this growth. However, we and others in the field also realize that behavioral science needs to evolve further over its next decade.

In this manifesto we take a clear-eyed look at the challenges facing the field and offer ten proposals for making further progress. As a starting point, we present the main arguments from critics of the behavioral insights approach on the following page.

THE CRITICISMS

Limited impact	Limited impact. The approach has focused on more tractable and easy-to-measure changes at the expense of bigger impact: it has just been tinkering around the edges of fundamental problems.[2]
Failure to reach scale	The approach promotes a model of experimentation followed by scaling, but it has not paid enough attention to how successful scaling happens - and the fact it often does not happen.[3]
Mechanistic thinking	The approach has promoted a simple, linear, and mechanistic way of understanding and influencing behavior that ignores second-order effects and spillovers (and employs evaluation methods that assume a move from A to B against a static background).[4]
Flawed evidence base	The replication crisis has challenged the evidence base underpinning the behavioral insights approach, adding to existing concerns like the duration of its interventions' effects.[5]
Lack of precision	The approach lacks the ability to construct precise interventions and establish what works for whom, and when. Instead, it relies either on over-general frameworks or disconnected lists of biases.[6]
Overconfidence	The approach is affected by the wider problem of over-confidence and can over-extrapolate from its evidence base, particularly when testing is not an option.[7]
Control paradigm	The approach can be elitist and pays insufficient attention to people's own goals and strategies; it uses concepts like "irrationality" to justify attempts to control the behavior of individuals, since they lack the means to do so themselves.[8]
Neglect of the social context	The approach has a limited, overly cognitive and individualistic view of behavior that neglects the reality that humans are embedded in established societies and practices.[9]
Ethical concerns	The behavioral insights approach will face more ethics, transparency, and privacy conundrums as it attempts more ambitious and innovative work.[10]
Homogeneity of participants and perspectives	The range of participants in behavioral science research has been narrow and unrepresentative; [11] homogeneity in the locations and personal characteristics of behavioral scientists influences their viewpoints, practices, and theories.[12]

THE PROPOSALS

We do not agree with all these criticisms, but we do think that they highlight several challenges that must - and can - be met. Doing so will mean behavioral science is better equipped to help build policies, products, and services on stronger empirical foundations - and thereby address the world's crucial challenges.

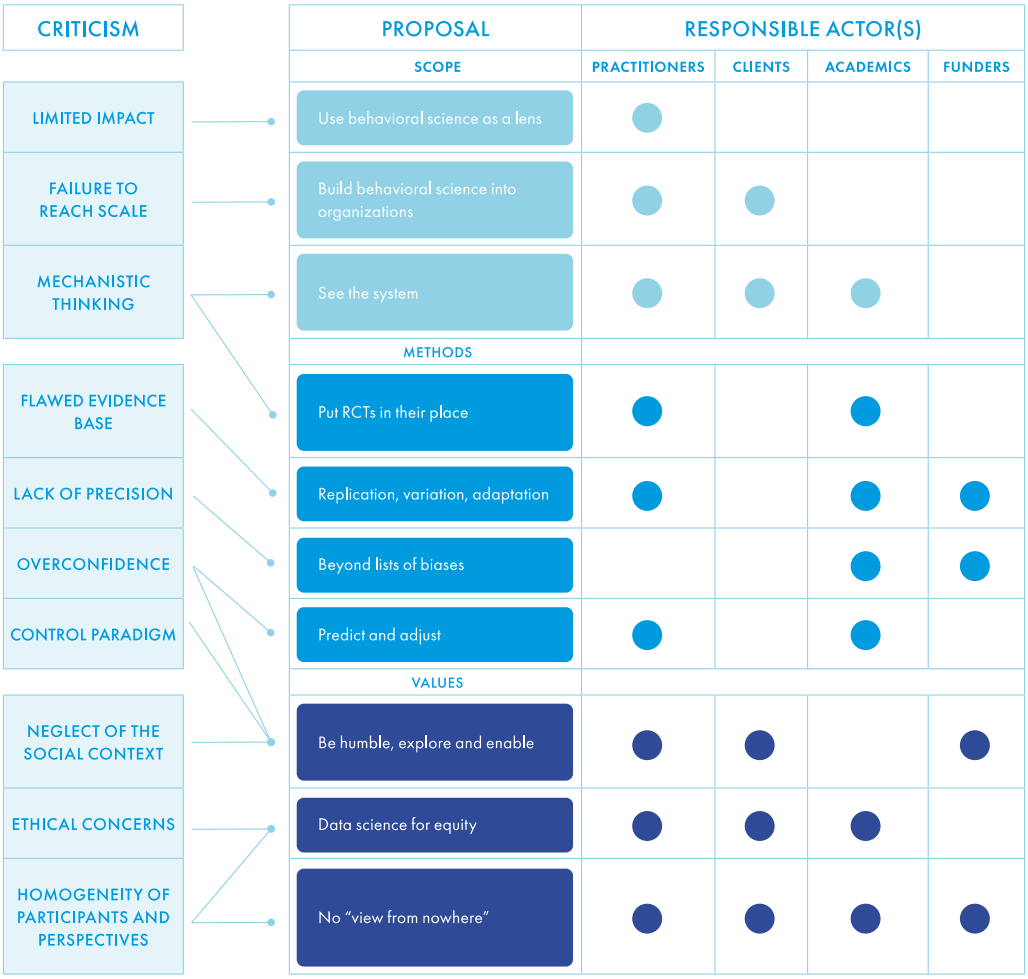
Our ten proposals for applied behavioral science fall into three categories: scope (the range and scale of issues to which behavioral science is applied); methods (the techniques and resources that behavioral science deploys); and values (the principles, ideals, and standards of conduct that behavioral scientists adopt).

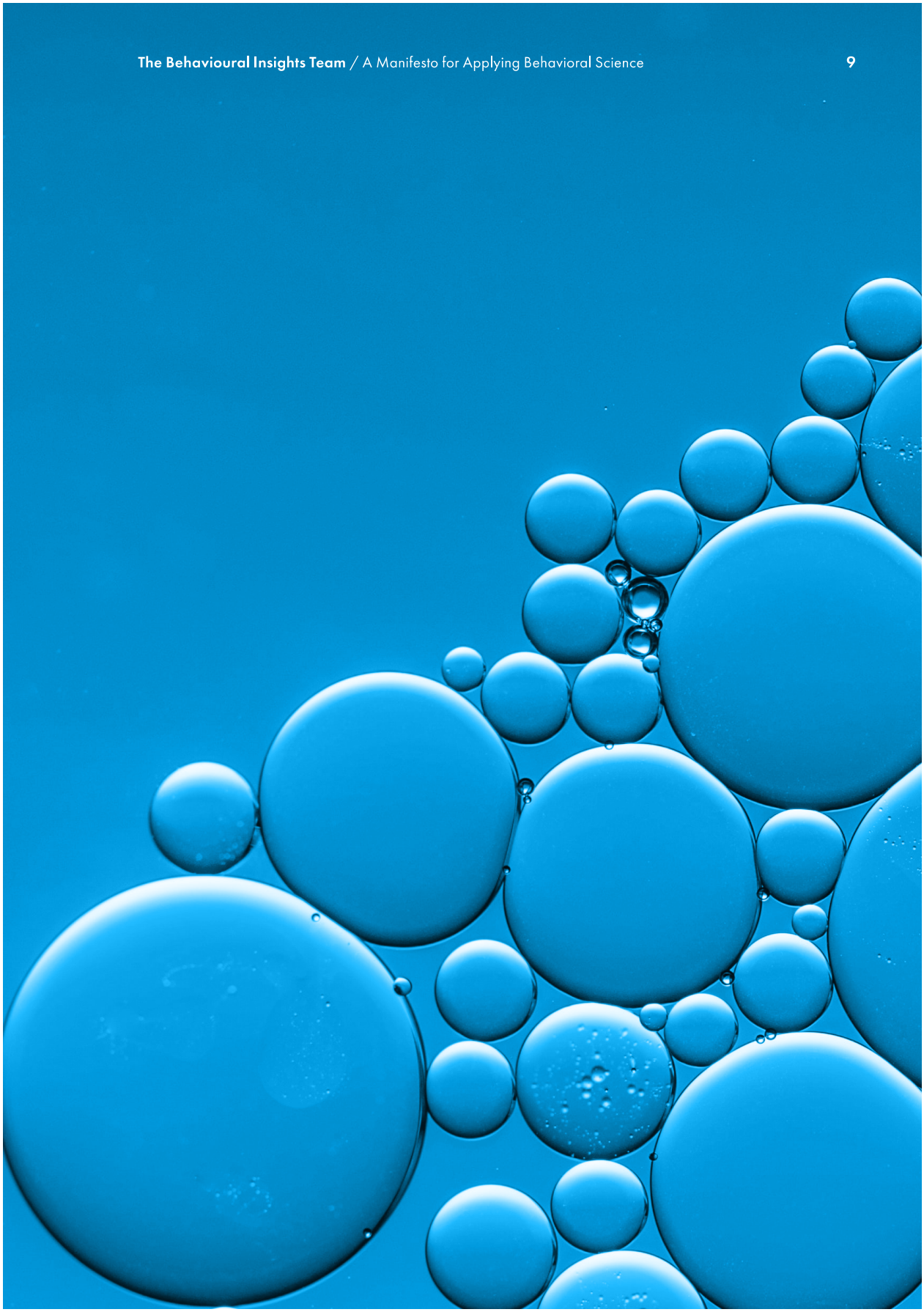
Category	Proposal	Recommended action(s)
Scope	01 **USE BEHAVIORAL SCIENCE AS A LENS** Summary: Page 11 Detail: Page 25	Present behavioral science as a lens that improves the view of any public and private issue, in order to break a self-sustaining pattern that has directed behavioral science away from the most significant problems.
	02 **BUILD BEHAVIORAL SCIENCE INTO ORGANIZATIONS** Summary: Page 12 Detail: Page 29	Focus less on how to set up a dedicated behavioral science team, and more on how the approach can be integrated into an organization's standard processes by upgrading its "choice infrastructure".
	03 **SEE THE SYSTEM** Summary: Page 14 Detail: Page 37	Use aspects of complexity thinking to improve behavioral science so it can: exploit "leverage points"; model the collective implications of heuristics; alter specific features of systems to create wider changes; and understand the longer-term impact on a system of a collection of policies with varying goals.

Category	Proposal	Recommended action(s)
Methods	04  **PUT RCTs IN THEIR PLACE** Summary: Page 15 Detail: Page 45	Strengthen RCTs to deal better with complexity by: gaining a better understanding of the system interactions and anticipate how they may play out; setting up RCTs to measure diffusion and contagion in networks; building feedback and adaptation into the design of RCTs and interventions.
	05  **REPLICATION, VARIATION AND ADAPTATION** Summary: Page 16 Detail: Page 51	Identify the most reliable interventions, develop an accurate sense of the likely size of their effects, and avoid the weaker options. Recognize that heterogeneity requires a much higher bar for claiming that an effect holds true across many unspecified settings. Create multi-site studies to systematically study heterogeneity in a wider range of contexts and participants. Codify and cultivate the practical skills that successfully adapt interventions to new contexts.
	06  **BEYOND LISTS OF BIASES** Summary: Page 17 Detail: Page 61	Emphasize theories that are "practical": they fill the gap between high-level frameworks and jumbled lists of biases; they are based on data and generate testable hypotheses, but also specify the conditions under which a prediction applies; they present actionable steps to solve real-world problems.
	07  **PREDICT AND ADJUST** Summary: Page 18 Detail: Page 69	Develop the practice of getting behavioral scientists to predict the results of experiments, and then feeding back the results to them.

Category	Proposal	Recommended action(s)
Values	08 **BE HUMBLE, EXPLORE AND ENABLE** Summary: Page 18 Detail: Page 75	Avoid using the term "irrationality"; practice "epistemic humility"; and design processes and institutions to counteract overconfidence. Pay greater attention to people's own interpretations of their beliefs, feelings and behaviors. Reach a wider range of experiences, including marginalized voices and communities. Recognize how apparently universal cognitive processes are shaped by specific contexts. Use six criteria (detailed in the main text) to assess when to enable people to use behavioral science themselves.
	09 **DATA SCIENCE FOR EQUITY** Summary: Page 20 Detail: Page 85	Use data science to identify the ways in which an intervention or situation appears to increase inequalities and introduce features to reduce them. For example, groups that are particularly likely to miss a filing requirement could be offered pre-emptive help.
	10 **NO VIEW FROM NOWHERE** Summary: Page 21 Detail: Page 91	Cultivate self-scrutiny; find new ways for the subjects of research to judge researchers; take actions to increase diversity among behavioral scientists and their teams, such as building professional networks between the Global North and Global South.

The figure below shows how each proposal maps onto the criticisms, as well as which groups have responsibility for implementing them: practitioners (individuals or teams who apply behavioral science findings in practical settings); the clients who commission these practitioners (for example, public or private sector organizations); academics working in the behavioral sciences (including disciplines such as anthropology, economics, and sociology); and funders who support the work of these academics.





SCOPE

01 USE BEHAVIORAL SCIENCE AS A LENS

The early phase of the behavioral insights movement was marked by skepticism about whether findings from laboratories would translate to real-world settings.[13] Mindful of this concern, practitioners developed standard approaches that could demonstrate a clear causal link between an intervention and an outcome.[14]

In practice, these approaches directed attention towards how the design of specific aspects of a policy, product or service influences discrete behaviors by actors who are considered mostly in isolation.[15] These standard approaches are strong and have produced compelling results. But they have also encouraged people to see behavioral science as a kind of specialist tool. This view mostly limits behavioral science to fixing concrete aspects of predetermined interventions - rather than shaping broader policy goals. Behavioral science acts as an alternative to standard tools, and it should be applied only to certain kinds of "behavioral" issues.[16]

Such a view is both misguided and profoundly limiting, but over time it has created a self-reinforcing perception that only certain kinds of tasks are "suitable" for behavioral scientists.[17] Opportunities, skills and ambitions have been constricted as a result.

A rebalancing is needed. Behavioral science also has much to say about pressing societal issues like discrimination, pollution, or economic mobility, and the structures that produce them.[18] These ambitions have always been present in the behavioral insights movement,[19] but the factors just outlined acted against them being realized more fully.[20]

The first step is to change the way we frame behavioral science itself. We need to see behavioral science as a lens that can be applied to any public and private issue. Using this frame shows that behavioral insights can enhance the way we see policy options (for example, revealing new ways of structuring taxes), rather than just acting as an alternative to them; it also conveys that creating new interventions to change behavior is not always the goal - which means more weight should be placed on the behavioral diagnosis of an issue. Behavioral science itself shows us the power of framing: the metaphors we use shape the way we behave, and therefore can be an agent of change.[21] Metaphors are particularly important in this case because the task of broadening the use of behavioral science requires making a compelling case to decision makers.[22] The metaphor of behavioral insights as a tool has established credibility and acceptance in a defined area; expanding beyond that area is the task for the next decade.

Full detail on page 25

SCOPE

02 BUILD BEHAVIORAL SCIENCE INTO ORGANIZATIONS

There has been too little focus on using behavioral science to shape organizations themselves, as opposed to increasing how much an organization uses behavioral science to achieve its goals.[23] We need to talk less on how to set up a dedicated behavioral team, and more about how behavioral science can be integrated into an organization's standard processes. For example, as well as trying to ensure that a departmental budget includes provisions for behavioral science, why not use behavioral science to improve the way this budget is created (e.g., are managers anchored to outdated spending assumptions)?

But we need to understand how this new way of thinking maps against existing debates about how to set up a behavioral function in organizations. We propose that doing so reveals six main scenarios, as shown in the diagram below.

In the "Baseline" scenario there is limited awareness of behavioral science in the organization, and its principles are not incorporated into processes. In the "Nudged Organization," levels of behavioral science awareness are still low, but its principles have been used to redesign processes to create better outcomes for staff or service users. No explicit behavioral science knowledge or capacity is created or needed, which means the return on investment here could be large. For that reason, this model feels like a neglected opportunity.

In "Proactive Consultancy", leaders may have set up a dedicated behavioral team without enough supporting organizational changes. The result is that the team has to work in an enterprising way, going to look for opportunities and having to prove its worth. But these teams may not be in a resilient position, since they lack ways to be grafted onto the standard processes of an organization.

Greater potential for scale ↓

		BEHAVIORAL SCIENCE KNOWLEDGE AND CAPACITY		
		LIMITED	CONCENTRATED	DIFFUSED
BEHAVIORAL SCIENCE INCORPORATED INTO ORGANIZATIONAL PROCESSES	NO	Baseline	Proactive consultancy	Behavioral entrepreneurs
	YES	Nudged organization	"Call for the experts"	Behaviorally-enabled organization

In "Call For The Experts", an organization has concentrated behavioral expertise, but there are also prompts and resources that allow this expertise to be integrated more into "business as usual". Expertise is not widespread, but access to it is. This setup could mean that processes stimulate demand for behavioral expertise that the central team can fulfill. That team may also have the institutional support to proactively monitor activities and respond quickly to specific crises.

In "Behavioral Entrepreneurs", there is behavioral science capacity distributed throughout the organization, either through direct capacity building or recruitment. The problem is that organizational processes do not support these individual pockets of knowledge. Therefore, those with expertise find it hard to apply ideas in practice, evaluate their effects, share findings, and build learning.

Finally, a "Behaviorally-Enabled Organization" is one where there is knowledge of behavioral science diffused throughout the organization, which also has processes that reflect this knowledge and support its deployment. Staff apply behavioral science in a deliberate way as part of "business as usual", rather than as special projects. While this is the most resilient setup, it also requires the most resources.

Most discussions make it seem like the meaningful choice is between the different columns in the table above - how to organize dedicated behavioral science resources. Instead, the more important move is from the top row to the bottom row: moving from projects to processes, from commissions to culture. A useful way of thinking about this task is about building or upgrading the "choice infrastructure" of the organization.[24]

Working out how best to build the choice infrastructure in organizations should be a major priority for behavioral science. One advantage to this approach is that it can help organizations address problems with scaling interventions. Already we can see some features will be crucial: reducing the costs of experimentation; creating a system that can learn from its actions; and developing new and better ways of using behavioral science principles to analyze the behavioral effects of organizational processes, rules, incentives, metrics, and guidelines.[25]

Full detail on page 29

SCOPE

03 SEE THE SYSTEM

Many important policy challenges emerge from complex adaptive systems, where change often does not happen in a linear or easily predictable way, and where coherent behavior can emerge from interactions without top-down direction.[26] There are many examples of such systems in human societies, including cities, markets, and political movements.[27] These systems can create "wicked problems" - like the Covid-19 pandemic - where ideas of success are contested, changes are non-linear and difficult to model, and policies have unintended consequences.[28]

This reality challenges the dominant behavioral science approach, which usually assumes stability over time, keeps a tight focus on predefined target behaviors, and predicts linear effects based on a predetermined theory of change.[29] The end result, some argue, is a failure to understand how actors are acting and reacting in a complex system that leads policymakers to conclude they are being "irrational" - and then actually disrupt the system in misguided attempts to correct perceived biases.[30] Behavioral science can be improved by using aspects of complexity thinking to offer new, credible, and practical ways of addressing major policy issues. First, we need to reject crude distinctions of "upstream" versus "downstream" or the "individual frame" versus the "system frame".[31] Instead, complex adaptive systems show that "higher-level" features of a system can actually emerge from the "lower-level" interactions of actors participating in the system.[32] When they become the governing features of the system, they then shape the "lower-level" behavior until some other aspect emerges, and the fluctuations continue. We can see this pattern in the way that new coronavirus variants emerged from specific contexts to re-shape the whole course of the pandemic.

In other words, we are dealing with 'cross-scale behaviors'.[33] For example, norms, rules, practices, and culture itself can emerge from aggregated social interactions; these features then shape cognition and behavioral patterns in turn.[34]

Recognizing cross-scale behaviors means that behavioral science could:

- Identify "leverage points'' where a specific shift in behavior will produce wider system effects. For example, if even a subset of consumers decides to switch to a healthier version of a food product, this can have broader effects on a population's consumption through the way the food system responds by restocking and product reformulation.[35]
- Model the collective implications of individuals using simple heuristics to navigate a system. For example, new models show how small changes to simple heuristics that guide savings (in this case, how quickly households copy the savings behaviors of neighbors) can lead to inequalities in wealth suddenly emerging.[36]
- Find targeted changes to features of a system that create the conditions for wide-ranging shifts in behavior to occur. For example, a core driver for social media behaviors is the ease with which information can be shared.[37] Even minor changes to this factor can drive widespread changes - some have argued that such a change is what created the conditions leading to the Arab Spring, for example.[38]

This approach also suggests that a broader change in perspective is needed. We need to realize the flaws in launching interventions in isolation and then moving on when a narrowly defined goal has been achieved. Instead, we need to see the longer-term impact on a system of a collection of policies with varying goals.[39] The best approach may be "system stewardship", which focuses on creating the conditions for behaviors and indirectly steering adaptation towards overall goals.[40]

Of course, not every problem will involve a complex adaptive system; for simple issues the standard behavioral approach works well. So behavioral scientists should develop the skills to recognize the type of system that they are facing ("see the system"), and then choose their approach accordingly. These skills can be developed through agent-based simulations,[41] immersive technologies,[42] or just basic checklists.[43]

Full detail on page 37

METHODS

04 PUT RCTs IN THEIR PLACE

Randomized Controlled Trials (RCTs) have been a core part of applied behavioral science, and they work very well in relatively simple and stable contexts. But they can fare worse in complex adaptive systems, whose many shifting connections can make it difficult to keep a control group isolated, and where a narrow focus on predetermined outcomes may neglect others that are important but difficult to predict.[44]

We can strengthen RCTs to deal better with complexity. We can try to gain a better understanding of the system interactions and anticipate how they may play out, perhaps through "dark logic" exercises that try to trace potential harms, rather than benefits.[45] We can set up RCTs to measure diffusion and contagion in networks, either by creating separate online environments or by randomizing real-world clusters, like separate villages.[46]

Finally, we can build feedback and adaptation into the RCT design, allowing adjustments to changing conditions.[47] Options include using two-stage trial protocols,[48] evolutionary RCTs,[49] sequential multiple assignment randomized (SMART) trials,[50] and "bandit" algorithms that identify high-performing interventions and allocate more people to them.[51]

We can also use behavioral science to enhance alternative ways of measuring impact - in particular, agent-based modeling, which tries to simulate the interactions between the different actors in a system.[52] The agents in these models are mostly assumed to be operating on rational choice principles.[53] Therefore, there is a big opportunity to build in more evidence about the drivers of behavior - for example, habits and social comparisons.[54]

Full detail on page 45

METHODS

05 REPLICATION, VARIATION, ADAPTATION

The "replication crisis" of the last decade has seen intense debate and concern about the reliability of behavioral science findings. Poor research practices were a major cause of the replication crisis; the good news is that many have improved as a result.[55]

We need to secure and build on these advances, so that we move towards a future where meta-analyses of high-quality studies (including deliberate replications) are used to identify the most reliable interventions, develop an accurate sense of the likely size of their effects, and avoid the weaker options. We have a responsibility to discard ideas if solid evidence now shows they are shaky, and to offer a realistic view of what behavioral science can accomplish.

That responsibility also requires us to have a hard conversation about heterogeneity in results: the complexity of human behavior creates so much statistical "noise" that it's often hard to detect consistent signals and patterns.[56] The main drivers of heterogeneity are that a) contexts influence results and b) the effect of an intervention may vary greatly between groups within a population.[57] These factors complicate the idea of replication itself: a "failed" replication may not show that a finding was false, but rather how it exists under some conditions and not others.[58]

These challenges mean that applied behavioral scientists need to set a much higher bar for claiming that an effect holds true across many unspecified settings.[59] There is a growing sense that interventions should be talked about as hypotheses that were true in one place, and which may need adapting for them to be true elsewhere as well.[60]

We need specific proposals as well as narrative changes. The first concerns data collection: expand studies to include (and thus examine) a wider range of contexts and participants, and gather richer data about them. To date, only a small minority of behavioral studies have provided enough information to see how effects vary.[61] Coordinated multi-site studies will be needed to collect enough data to explore heterogeneity systematically; "crowdsourced" studies offer particular promise for testing context and methods.[62]

Behavioral scientists also need to get better at judging how much an intervention's results were linked to its context - and therefore how much adaptation it may need.[63] We should use and modify frameworks from implementation science to develop such judgment.[64] Finally, we need to codify and cultivate the practical skills that successfully adapt interventions to new contexts; expertise in behavioral science should not be seen as simply knowing about concepts and findings in the abstract. Therefore, it's particularly valuable to learn from practitioners how they adapted specific interventions to new contexts. These accounts are starting to emerge, but they are still rare,[65] since researchers are incentivized to claim universality for their results, rather than report and value contextual details.[66]

Full detail on page 51

METHODS

06 BEYOND LISTS OF BIASES

The heterogeneity in behavioral science findings also means that our underlying theories need to improve: we are lacking good explanations for why findings vary so much.[67] This need for better theories can be seen as part of a wider "theory crisis" in psychology, which has thrown up two big concerns for behavioral science.[68]

The first stems from the fact that theories of behaviour often try to explain phenomena that are complex and wide-ranging.[69] Trying to cover this variability can produce descriptions of relationships and definitions of constructs that are abstract and imprecise. The result is theories that are vague and "weak", since they can be used to generate many different hypotheses - some of which may actually contradict each other.[70] That makes theories hard to disprove and so weak theories stumble on, unimproved.[71]

The other concern is that theories can make specific predictions, but they are disconnected from each other - and from a deeper, general framework that can provide broader explanations (like evolutionary theory, for example).[72] The main way this issue affects behavioral science is through heuristics and biases. Examples of individual biases are accessible, popular, and how many people first encounter behavioral science. These ideas are incredibly useful, but have often been presented as lists of standalone curiosities, in a way that is incoherent, reductive, and deadening. They can create overconfident thinking that targeting a specific bias (in isolation) will achieve a certain outcome.[73]

Perhaps most importantly, focusing on lists of biases distracts us from answering core underlying questions. When does one or another bias apply? Which are widely applicable, and which are highly specific? These are highly practical questions when someone is faced with tasks like, for example, taking an intervention to new places.

The concern for behavioral science is that it uses both these high-level frameworks, like dual process theories, and jumbled collections of heuristics and biases - with little in the middle to draw both levels together.[74]

We think that a priority for responding to this challenge is to develop theories that are practical. By this we mean:

- They fill the gap we've identified in behavioral science: between day-to-day working hypotheses and comprehensive and systematic attempts to find universal underlying explanations.
- They are based on data rather than being derived from pure theorizing.[75]
- They can generate testable hypotheses, so they can be disproved.[76]
- However, they also specify the conditions under which a prediction applies or does not.[77]
- They are geared towards realistic adaptation by practitioners and offer 'actionable steps toward solving a problem that currently exists in a particular context in the real world.'[78]

We think that resource rationality is a good example of a practical theory. It starts from the basis that people make rational use of their limited cognitive resources.[79] Given there is a cost to thinking, people will look for solutions that balance choice quality with effort. Importantly, these principles offer a systematic framework for building useful models for how people act.

A recent study has shown how these models can can not only predict how people will respond to different kinds of nudges in certain contexts, but also can be integrated with machine learning to create an automated method for constructing 'optimal nudges'.[80] These are highly practical benefits coming from applying a particular theory.

Full detail on page 61

METHODS

07 PREDICT AND ADJUST

Hindsight bias is what happens when people feel "I knew it all along", even if they did not.[81] When the results of an experiment come in, hindsight bias may mean that behavioral scientists are more likely to think that they had predicted them, or quickly find ways of explaining why they occurred. Hindsight bias is a big problem because it breeds overconfidence, impedes learning, dissuades innovation, and prevents us from understanding what is truly unexpected.[82]

In response, behavioral scientists should establish a standard practice of predicting the results of experiments, and then receiving feedback on how their predictions performed. Hindsight bias can flourish if we do not systematically capture expectations or "priors" about what the results of a study will be - in other words, it is not easy to check or remember the state of knowledge before an experiment.[83] Making predictions provides regular, clear feedback of the kind that is more likely to trigger surprise and reassessment, rather than hindsight bias.[84]

More and more studies are explicitly integrating predictions.[85] But barriers lie in the way of further progress. People may not welcome the ensuing challenge to their self-image; predicting may seem like one thing too many on the to-do list; and the benefits lie in the future.

We propose: make predicting easy by incorporating it into standard organizational processes; minimize threats to predictors' self-image, for example by making and feeding back predictions anonymously;[86] give concrete prompts for learning and reflection, in order to disrupt the move from surprise to hindsight bias;[87] and build learning from prediction within and between institutions.

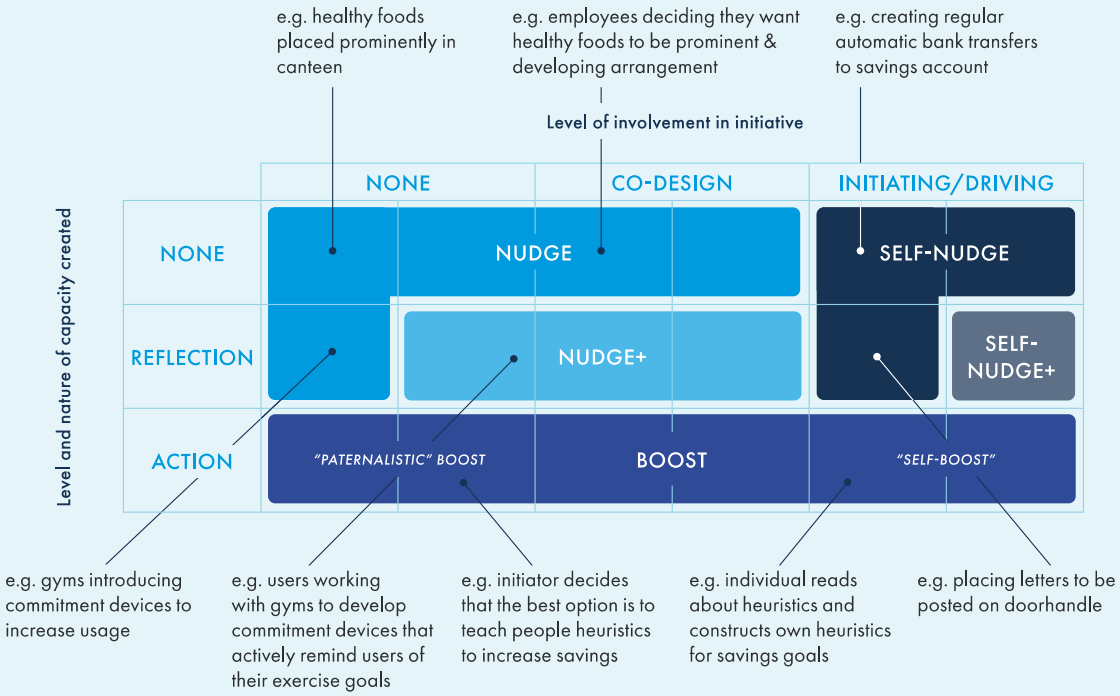
Full detail on page 69

VALUES

08 BE HUMBLE, EXPLORE AND ENABLE

Behavioral scientists (like other experts) may over-confidently rely on decontextualized principles that do not match the real-world setting for a behavior.[88] Deeper inquiry can reveal reasonable explanations for what seem to be behavioral biases.[89] In response, those applying behavioral science should: avoid using the term "irrationality", which can limit attempts to understand actions in context; acknowledge that diagnoses of behavior are provisional and incomplete ("epistemic humility");[90] and design processes and institutions to counteract overconfidence.[91]

A common theme through these ideas is the need for more and better inquiry into behaviors in context, rather than making assumptions. Open-ended qualitative exploration of the context and drivers for behaviors is not new to the behavioral sciences.[92] However, three areas demand particular focus in the future. First, pay greater attention to people's goals and strategies, and their own interpretations of their beliefs, feelings, and behaviors.[93] Second, reach a wider range of experiences, including marginalized voices and communities, understanding how structural inequalities can lead to expectations and experiences varying greatly by group and geography.[94] Third, recognize how apparently universal cognitive processes are shaped by specific contexts, thereby unlocking new ways for behavioral science to engage with values and culture.[95]



In addition, more can and should be done to broaden ownership of behavioral science approaches. Many, but far from all, behavioral science applications have been quite top-down, with a 'choice architect' enabling certain outcomes.[96] One route is to enable people to become more involved in designing interventions themselves - and "nudge plus", "self nudges", and "boosts" have been proposed as ways of doing this.[97] Reliable criteria are needed to decide when enabling approaches may be appropriate, including: whether the opportunity to use an enabling approach exists; ability and motivation; preferences; learning and setup costs; equity impacts; and effectiveness (recognizing evidence on this point is still emerging).[98]

But these new approaches should not be seen simplistically as "enabling" alternatives to "disempowering" nudges.[99] Instead, we need to consider a) how far the person performing the behavior is involved in shaping the initiative itself; b) the level and nature of any capacity created by the intervention.

People may be heavily engaged in selecting and developing a nudge intervention that nonetheless does not trigger any reflection or build any skills. Alternatively, a policy maker may have paternalistically assumed that people want to build up their capacity to perform an action, when in fact they do not. This is the real choice to be made.

A final piece missing from current thinking is that enabling people can lead to a major decentering of the use of behavioral science. If more people are enabled to use behavioral science, they may decide to introduce interventions that influence others. Rather than just creating self-nudges through altering their immediate environments, they may decide that wider system changes are needed instead. A range of people could be enabled to create nudges that generate positive societal change (with no "central" actors involved), as happened for the "Fair Tax Mark" in the UK.

Full detail on page 75

VALUES

09 DATA SCIENCE FOR EQUITY

Recent years have seen growing interest in using new data science techniques to reliably analyze the heterogeneity of large datasets.[100] Machine learning is claimed to offer more sophisticated, reliable, and data-driven ways of detecting meaningful patterns in datasets.[101] For example, a machine learning approach has been shown to be more effective than conventional segmentation approaches at analyzing patterns of US household energy usage to reduce peak consumption.[102]

A popular idea is to use such techniques to better understand what works best for certain groups, and thereby tailor an offering to them.[103] "Scaling" an intervention stops being about a uniform roll-out, and instead becomes about presenting recipients with the aspects that are most effective for them.[104]

This vision is often presented as straightforward and obviously desirable, but it runs almost immediately into ethical quandaries and value judgements. People are unlikely to know what data has been used to target them, and how; the specificity of the data involved may make manipulation more likely, since it may exploit sensitive personal vulnerabilities; and expectations of universality and non-discrimination in public services may be violated.[105]

There is also emerging evidence that people often object to personalization. While they support some personalized services, they consistently oppose advertising that is customized based on sensitive information - and they are generally against the collection of the information that personalization relies on.[106] When a company tries personalization that crosses into being "creepy," uproar and damage to its reputation can ensue.[107]

In order to navigate this landscape, behavioral scientists need to examine four factors.

- **Who** does the personalization target, and using what criteria? Many places have laws or norms to ensure equal treatment based on personal characteristics. When does personalization violate those principles?
- **How** is the intervention constructed? To what extent do the recipients have awareness of the personalization, choice over whether it occurs, control over its level or nature, and the opportunity for giving feedback on it?[108]
- **When** is it directed? Is it at a time when the participant is vulnerable? Would they likely regret it later, if they had time to reflect?
- **Why** is personalization happening? Does it aim to exploit and harm or support and protect, recognizing that those terms are often contested?

Taking these factors into account, we propose that the main opportunity is for data science to identify the ways in which an intervention or situation appears to increase inequalities, and reduce them.[109] For example, groups that are particularly likely to, say, miss a filing requirement, could be offered preemptive help.

We call this idea data science for equity. It addresses the "why" factor by using data science to support not exploit. But it needs to be supported by other attempts to increase agency (the "how" factors), like a recent study that showed how boosts can be used to help people detect micro-targeting of advertising,[110] and studies that obtain more data on which uses of personalization people find acceptable.

Full detail on page 85

VALUES

10 NO VIEW FROM NOWHERE

Behavioral scientists need to understand how they bring certain assumptions, privileges, and ways of seeing to what they do.[111] They are always situated, embedded, and entangled with ideas and situations. They cannot assume there is some set-aside position from which to observe the behavior of others - there is no "view from nowhere".[112]

Behavioral scientists are defined by having knowledge, skills, and education; many of them can use these resources to shape public and private actions. Therefore, they are in a privileged position, but may not see the extent to which they hold elite positions that stop them from understanding people who think differently (for example those who are skeptical of education).[113]

There have been repeated concerns that the field is still highly homogeneous in other ways as well. Gender, race, physical abilities, sexuality, and geography also influence the viewpoints, practices, and theories of behavioral scientists.[114] Only a quarter of the behavioural insights teams catalogued in a 2020 survey were based in the Global South.[115] The last decade has shown just how behaviors can vary greatly from culture to culture, even as psychology has tended to generalize from relatively small and unrepresentative samples.[116] So, rather than claiming that science is value-free, we need to find realistic ways of acknowledging and improving this reality.[117]

A starting point is for behavioral scientists to cultivate self-scrutiny by querying how their identities and experiences contribute to their stance on a topic. Hypothesis generation could particularly benefit from this exercise, since arguably it is closely informed by the researcher's personal priorities and preferences.[118] Behavioral scientists could be actively reflecting on interventions in progress, including what factors are contributing to power dynamics.[119]

Self-scrutiny may not be enough. We should also find ways for people to judge researchers and decide whether they want to participate in research - going beyond consent forms. Finally, we should take actions to increase diversity (of several kinds) among behavioral scientists, teams, collaborations, and institutions. These could include increased support for starting and completing PhDs, reducing the significant racial gaps present in much public funding of research, or building professional networks that connect the Global North and Global South.[120]

Full detail on page 91

CONCLUSION

When considered together, these proposals present a consistent and coherent vision for the future of applied behavioral science. A common theme throughout the ten proposals is the need for self-reflective practice. In other words, a main priority for behavioral scientists is to recognize the various ways that their own behavior is being shaped by structural, institutional, environmental, and cognitive factors.

However, as the field itself shows, a gap often emerges between intention and action. Given what's at stake, BIT will focus on bridging this gap in the coming years. Realizing these proposals will require sustained work and experiencing the discomfort of disrupting what may have become familiar and comfortable practices. Indeed, this manifesto forms part of a new collection of resources from BIT to start to fulfill the goals set out here.

Improving applied behavioral science has some characteristics of a social dilemma - benefits are diffused across the field as a whole, while costs fall on any individual party who chooses to act (or act first). Practitioners are often in competition. Academics often want to establish a distinctive research agenda. Commissioners are often rewarded for risk aversion. Impaired coordination is particularly problematic, since it forms the basis for several necessary actions (such as the multi-site studies to measure heterogeneity).

Solving these problems will be hard. Funders need to find mechanisms that adequately reward coordination and collaboration by recognizing the true costs involved. Practitioners need to perceive the competitive advantage from adopting new practices and be able to communicate them to clients. Stepping back, the starting point for these changes needs to be a change in the narrative about what the field does and could do. The "manifesto" presented here aims to help shape this narrative.

Category	Proposal	Recommended action(s)
Scope	**Use behavioral science as a lens**	Present behavioral science as a lens that improves the view of any public and private issue, in order to break a self-sustaining pattern that has directed behavioral science away from the most significant problems.
	Build behavioral science into organizations	Focus less on how to set up a dedicated behavioral science team, and more on how the approach can be integrated into an organization's standard processes by upgrading its "choice infrastructure".
	See the system	Use aspects of complexity thinking to improve behavioral science so it can: exploit "leverage points"; model the collective implications of heuristics; alter specific features of systems to create wider changes; and understand the longer-term impact on a system of a collection of policies with varying goals.

Category	Proposal	Recommended action(s)
Methods	Put RCTs in their place	Strengthen RCTs to deal better with complexity by: gaining a better understanding of the system interactions and anticipate how they may play out; setting up RCTs to measure diffusion and contagion in networks; building feedback and adaptation into the design of RCTs and interventions.
	Replication, variation, adaptation	Identify the most reliable interventions, develop an accurate sense of the likely size of their effects, and avoid the weaker options. Recognize that heterogeneity requires a much higher bar for claiming that an effect holds true across many unspecified settings. Create multi-site studies to systematically study heterogeneity in a wider range of contexts and participants. Codify and cultivate the practical skills that successfully adapt interventions to new contexts.
	Beyond lists of biases	Emphasize theories that are "practical": they fill the gap between high-level frameworks and jumbled lists of biases; they are based on data and generate testable hypotheses, but also specify the conditions under which a prediction applies; they present actionable steps to solve real-world problems.
	Predict and adjust	Develop the practice of getting behavioral scientists to predict the results of experiments, and then feeding back the results to them.
Values	Be humble, explore, and enable	Avoid using the term "irrationality"; practice "epistemic humility"; and design processes and institutions to counteract overconfidence. Pay greater attention to people's own interpretations of their beliefs, feelings, and behaviors. Reach a wider range of experiences, including marginalized voices and communities. Recognize how apparently universal cognitive processes are shaped by specific contexts. Use six criteria to assess when to enable people to use behavioral science themselves.
	Data science for equity	Use data science to identify the ways in which an intervention or situation appears to increase inequalities and introduce features to reduce them. For example, groups that are particularly likely to miss a filing requirement could be offered pre-emptive help.
	No "view from nowhere"	Cultivate self-scrutiny; find new ways for the subjects of research to judge researchers; take actions to increase diversity among behavioral scientists and their teams, such as building professional networks between the Global North and Global South.

SCOPE

01 USE BEHAVIORAL SCIENCE AS A LENS

> We need to see behavioral science as a lens that improves the view of any public and private issue, rather than as a tool that we sometimes pick up. Making this change will help break the self-sustaining pattern whereby demand for behavioral science, and the tools we have developed, has pushed work towards 'downstream' interventions and away from structural changes.

The recent surge in applying behavioral science to practical issues has made a measurable difference across many domains. The approach has been adopted by public sector bodies at the local, national, and supra-national level,[121] and by private companies large and small. They have improved outcomes in health,[122] education,[123] sustainability,[124] transport,[125] diet,[126] and financial behavior,[127] among many areas. Many of these improvements have come at relatively low cost.[128]

Despite these achievements, objections have emerged. A common one is that there's been a focus on tractable and easy-to-measure changes, at the expense of bigger impact on major issues. Behavioral science, it's claimed, has just been tinkering around the edges of fundamental problems.[129]

We agree with the challenge that behavioral science can and should do more. Every day, new policies cut against well-established evidence of how people behave.[130] Services are shaped in ways that people cannot navigate. Products are launched with fundamental misconceptions about how people are likely to approach them. There are fewer prominent examples that clearly show how governing policies and systems have been designed using concepts from behavioral science, as opposed to specific aspects of how those policies were presented or structured.[131]

So, how to move forward? Step one is to realize that the strengths that have brought success may also be holding behavioral science back.
To explain, let's go back to the start of the current phase of applied behavioral science (around 2008-2012), when there was a pressing need to demonstrate clear results and build credibility. That pressure led us and others to form standard ways of applying "behavioral insights". These approaches generally have a common set of features.[132] The standard series of actions looks something like this:

- scoping the issue and exploring drivers of behavior
- defining a specific target behavior that can be measured reliably
- generating evidence-based interventions to change the target behavior
- creating a robust experimental design to test the intervention's effects
- if desired, taking the intervention to new places (e.g., "scaling")

These actions are usually presented in a step-by-step (linear) way, although most guides stress that people can loop between stages.

Over the past decade, this kind of approach has tended to produce:

- "downstream" interventions that concern how specific aspects of a policy, product or service are designed
- a focus on discrete behaviors by actors (e.g., people, businesses), considered mostly in isolation.[133]

Perhaps a good example is BIT's project to reduce missed hospital appointments in the UK.[134] This work identified the wording of text message reminders as an opportunity for improvement. We ran two randomized controlled trials in London, which found that a message referring to the cost of a missed appointment for the health system reduced no-shows from 11.1% to 8.5% - a 25% relative change. This low-cost change was then taken up by other health providers around the world.[135]

As this example shows, the approach is a strong one. There's a neatly-defined problem, a specific intervention, and a strong causal link between that intervention and the target outcome. There's a clear story to tell about what happened and why. These strengths mean that there are still so many improvements that this approach could achieve. For instance, one priority should be to clear the vast administrative burdens (or 'sludge') that prevent people - particularly those with fewest resources - from understanding or accessing government services.[136]

However, the justified focus on these clear and credible results about downstream impact also becomes self-reinforcing. People start to think that this is the sole way that behavioral science can be applied. In turn, this perception shapes demand: only certain kinds of problems are seen as ones 'suitable' for behavioral scientists.[137]

Opportunities, skills, and ambitions have been constricted as a result. In general, practitioners have focused more on expanding the application of some "tried and tested" interventions to new areas, and less on exploring new ones (or getting a deeper understanding of familiar ones).[138]

We need a rebalancing. Behavioral science also has much to say about broader, larger issues in society like discrimination, pollution, or economic mobility, and the structures that produce them. Behavioral science has the potential to fundamentally change how we understand the factors shaping behavior and therefore how we constitute an issue and what is possible.[139]

Take the economy: behavioral science can show how to regulate markets differently;[140] how to design taxes to drive wider behavioral changes;[141] and even offer a vision for the future "behavioral economy" as a whole.[142]

As this list shows, there are examples of how behavioral science has tackled more complex, structural, "upstream" issues. But these examples are harder to communicate because they often deal with the fluid, murky, and fractured narratives of politics and policy-making. Unlike the clear, linear stories of the approach outlined above, the contribution of behavioral science may be difficult to trace or may play out over a long timeframe. It's easier to talk about the neat narratives instead, particularly since many people are aware of and curious about the idea of 'nudging', which is often associated (inaccurately) with small presentational tweaks only.[143]

The wide potential scope of applied behavioral science is an idea that BIT has promoted consistently since its creation.[144] But the self-reinforcing limiting factors we outlined have proved strong. Now the increasingly urgent question is: how can we successfully change behavioral science itself?

Our answer is to consider the proposals in this manifesto, which aim to create a package that can achieve that change. We can start by trying to switch the metaphors or frames through which we perceive behavioral science itself.

Behavioral science should be understood as a lens that enhances the view of any public and private issue, rather than as a tool that we sometimes pick up.

The trends we highlight above have tended to reinforce this tool metaphor, which encourages this way of thinking:

Behavioral science is a specialist tool that is applied to certain kinds of problems - and not others. Often these are defined "delivery" issues ("How do we structure this message?"), but sometimes it can be used to solve a "behavioral" problem as an alternative to more traditional approaches like rules and incentives.

These implications lead us down the wrong path. Instead, behavioral science should be understood as a lens that can be applied to any public and private action. This change offers several advantages:

- A lens metaphor shows that behavioral insights can enhance the way we see policy options (for example, revealing new ways of structuring taxes), rather than just acting as an alternative to them.
- A lens metaphor conveys that the uses of behavioral insights are not limited to creating new interventions. A behavioral science lens can, for example, help reassess existing actions and understand how they may have unintended effects. It emphasizes the behavioral diagnosis of a situation or issue, rather than pushing too soon to define a precise target outcome and intervention.[145]

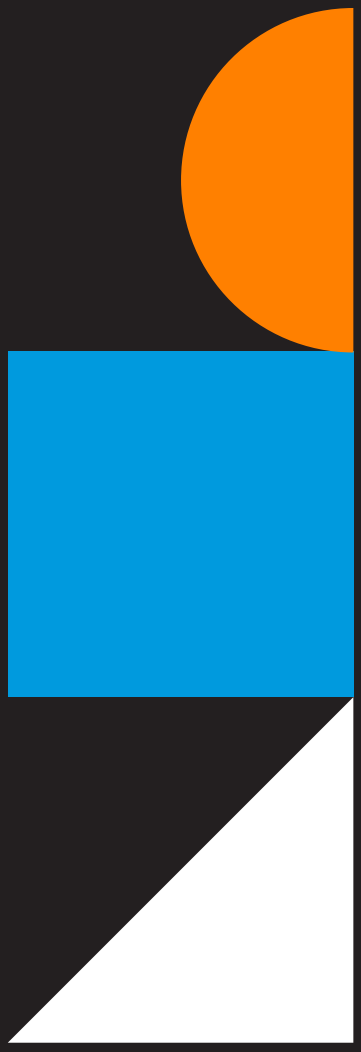
- Specifying that this lens can be applied to any action conveys the error of separating out "behavioral" and "non-behavioral" issues: most of the goals of private and public action depend on certain behaviors happening (or not). Behavioral science should therefore be integrated into an organization's core activities, rather than acting as an optional specialist tool.[146]

In one sense, this proposal is about returning to first principles. Back in 2010 we emphasized that 'civil servants...need to better understand the behavioral dimension of their policies and actions', and also stressed how behavioral science 'powerfully complements and improves conventional policy tools'.[147] But, for all the reasons above, this aspect has been less prominent over the last decade.[148]

Other metaphors apart from a lens would be powerful as well. For example, moving from 'choice architecture' to 'choice infrastructure' effectively highlights the broader, embedded nature of our behaviors.[149] The point is that behavioral science itself shows us the power of framing: the metaphors we use shape the way we behave, and therefore can be agents of change.[150]

Metaphors are particularly important because the task of broadening the use of behavioral science requires making a compelling case to decision makers. Behavioral science practitioners need to understand their audience and then shape their offers accordingly.

In a way, the metaphor of behavioral science as a tool that produces clear results has served the field well over the last decade - it has established credibility and acceptance in a defined area. The challenge now is to expand beyond that area, allowing behavioral science to fulfill its potential before the self-reinforcing cycle becomes too hard to break.



02

BUILD BEHAVIORAL SCIENCE INTO ORGANIZATIONS

> There has been too little focus on using behavioral science to shape organizations themselves, as opposed to increasing how much an organization uses behavioral science to achieve its goals. We need to talk less on how to set up a dedicated behavioral function, and more about how behavioral science can be integrated into an organization's standard processes.

For example, as well as trying to ensure that a departmental budget includes provisions for behavioral science, why not use behavioral science to improve the way this budget is created (e.g., are managers anchored to outdated spending assumptions)?

But we need to understand how this new way of thinking maps against the existing debate about how to set up a behavioral function in organizations. We propose that doing so reveals six main scenarios.

- In the **"Baseline"** scenario there is limited awareness of behavioral science in the organization, and its principles are not incorporated into processes.
- In the **"Nudged Organization,"** levels of behavioral science awareness are still low, but its principles have been used to redesign processes to create better outcomes for staff or service users.
- In **"Proactive Consultancy,"** leaders may have set up a dedicated behavioral team without grafting it onto the organization's standard processes. This lack of institutional grounding puts the team in a less resilient position, meaning it must always search for new work.
- In **"Call For Experts,"** an organization has concentrated behavioral expertise, but there are also prompts and resources that allow this expertise to be integrated more into "business as usual". Expertise is not widespread, but access to it is. This setup could mean that processes stimulate demand for behavioral expertise that the central team can fulfill.
- In **"Behavioral Entrepreneurs,"** there is behavioral science capacity distributed throughout the organization, either through direct capacity building or recruitment. The problem is that organizational processes do not support these individual pockets of knowledge.
- Finally, a **"Behaviorally-Enabled Organization"** is one where there is knowledge of behavioral science diffused throughout the organization, which also has processes that reflect this knowledge and support its deployment.

The common success factor in these scenarios is an upgrade of the "choice infrastructure" of organizations. To do this, we propose: reducing the costs of experimentation, creating a system that can learn from its actions; and developing new and better ways of using behavioral science knowledge to analyze the behavioral effects of processes, rules, incentives, metrics, and guidelines.

The second proposal is to broaden the scope of how behavioral science is used in organizations. Much attention has been paid to how the practice of applying behavioral science can have more influence within organizations – usually by advising on how a dedicated behavioral science function should be structured.[151]

This is an important question: there were more than 300 such units by 2020 worldwide; BIT and others have advised on setting them up.[152] Work in this area has covered important questions like how to arrange the leadership, organizational structure, funding, and goals of behavioral insights teams for maximum success.[153]

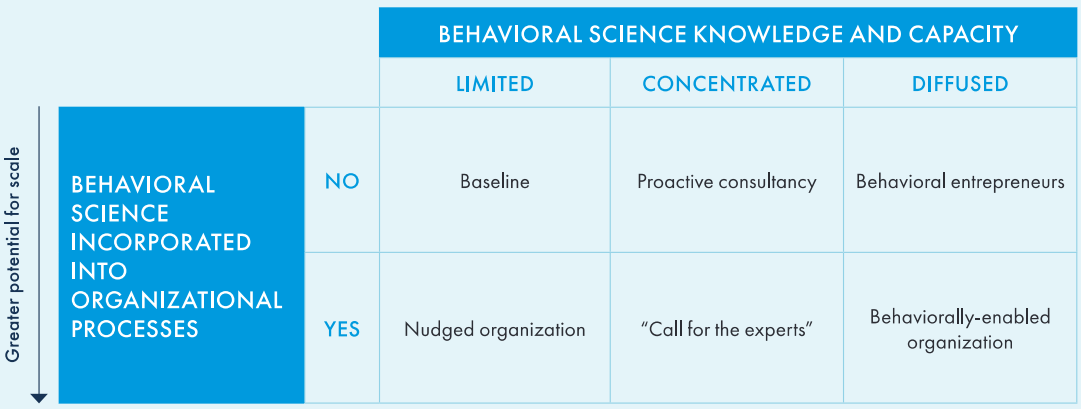
In contrast, there has been less attention paid to how behavioral science can be integrated into an organization's own processes.[154] There has not been enough focus on using behavioral science to shape organizations themselves, as opposed to increasing how much an organization uses behavioral science to achieve its goals.

For example, as well as trying to ensure that a departmental budget includes provisions for behavioral science, why not use behavioral science to improve the way this budget is created (e.g., are managers anchored to outdated spending assumptions)?[155]

The overriding message here is for greater focus on the organizational changes that indirectly apply or support behavioral science principles, rather than just thinking through how the direct and overt use of behavioral science can be promoted in an organization. There are two main advantages to doing this:

SCALE

Building behavioral science into organizations can address some of the issues surrounding successful scaling of interventions.[156] If some of the barriers to scaling concern cognitive biases in organizations, these changes could minimize the effect of such biases.[157] Rather than starting with a behavioral science project and then trying to scale it, we could start by looking at operations at scale and understand how they can be influenced. There is a change of perspective here to a position where 'what is scalable is not the content of what is learned in any given context, but the capacity for learning itself'.[158]



RESILIENCE

The goal would be to produce behaviorally-informed standard or "business as usual" processes, rather than the continued application of behavioral science explicitly. That approach is resilient to changes in the demand for "behavioral science" solutions as such in the future.

BIT's 2018 Behavioral Government report proposes many practical changes to organizational processes to mitigate biases in government.[159] But we need to understand the range of options for implementing such changes. How do we think about them alongside the desire to create a dedicated behavioral insights team, for example?

We think that the diagram above offers a useful way of mapping the options for building behavioral science into organizations.[160]

The vertical axis represents whether behavioral science has been used to shape the organization's own structures or processes, using a crude yes/no distinction to make the diagram manageable. We will bring this distinction to life with examples in the following sections, before defining it in more detail.

The horizontal axis deals with the extent and form of behavioral science knowledge and capacity in an organization.[161] In the 'Baseline' scenario, there is very little or no awareness of behavioral science concepts in the organization. 'Concentrated' refers to the setup where there is a dedicated team or resource that applies behavioral science to organizational priorities.[162] In the 'Diffused' scenario, people or teams with competence in behavioral science are spread throughout the organization. Deciding between concentrated and diffused setups is generally seen as a central choice for organizations looking to build a behavioral science function.[163]

We can now see how this framework illuminates the different options:

BASELINE

Here, there is limited awareness of behavioral science in the organization, and its principles are not incorporated into processes. Benefits are likely to be limited.

NUDGED ORGANIZATION

Here, levels of behavioral science awareness are still low, but its principles have been used to redesign processes to create better outcomes for staff or service users. For example, the pervasive optimism bias in organizations' plans can be reduced by mandatory 'pre-mortems', where decision makers imagine the future failure of their project and then work back to identify why things went wrong. Group reinforcement (or "groupthink") could be minimized by creating routes for diverse views to be fed in anonymously before and after group discussions, counteracting the pressure to conform face to face.[164] And there are vast opportunities to reduce the administrative burdens, or 'sludge', that make services and processes difficult to access and navigate.[165]

This scenario is termed the 'nudged organization' because no explicit behavioral science knowledge or capacity is created or needed. Like for nudging, it is the choice architecture (or choice infrastructure) that produces the outcomes, and there is no neutral choice in the way that an organization's processes are set up. That means no behavioral team or unit is created; the change or goals may not even be framed in terms of behavioral science (as for "administrative burdens")

For this reason, the best starting point is to understand how the existing setup is influencing behavior. Where is the choice architecture currently working well, through accident or design? How can existing processes be amended easily to draw on these practices? Who are the people who oversee the rules, incentives, metrics, and guidelines that influence people throughout the organization?

To give a concrete example, human resource leaders profoundly shape what organizations permit and reward. Yet, there has been relatively little focus on "behavioral HR". Recent studies have shown that cognitive biases such as decoy effects, framing effects, anchoring and halo effects can be created in practical decisions such as procurement and performance appraisal.[166] They can also be countered: when considering the purchase of an email software, framing effects like saying 20% of users were dissatisfied significantly affected intentions (versus saying 80% were satisfied), but these effects were eliminated if both percentages were shown (in a random order).[167]

The big outstanding question in this scenario is who introduces the nudges, since the organization has little internal capacity. Perhaps these could be one-off changes introduced from outside? Answering this question feels important, since the return on investment here could be large - and, for that reason, this model feels like a neglected opportunity that needs more attention.

PROACTIVE CONSULTANCY

In this situation, leaders may have set up a dedicated behavioral team, but perhaps not given much thought to supportive organizational changes. The result is that the team has to work in an enterprising way, going to look for opportunities and having to prove its worth.[168]

This situation reflects the reality for many teams, who are 'looking to develop networks, positions, and tactics that establish their authority and credibility among decision makers.'[169] As a result, much of the discussion has focused on how best to set up these teams. The better contributions have recognized that this question is fundamentally political, rather than technocratic - how do the people leading such a resource build relationships and present their team as useful to their organizations?

The problem with this scenario is that teams may not be in a resilient position, since they lack ways to be grafted onto the standard processes of an organization. For example, leaders may neglect to support and resource evidence-gathering and experimentation. At the same time, they may have unrealistic expectations because they know only the highlights of previous behavioral science success stories.[170]

Teams will therefore have to continually prove their worth and relevance, while having fewer options for doing so. As current practitioners point out, 'interventions that seem relatively easy to implement (e.g., an RCT with promotion letters) can require a set of system changes that touch a variety of groups in the organization (e.g., printing services, database administrators, processing centers).'[171] This kind of broader organizational scaffolding is not always prioritized, despite being needed for behavioral teams to fulfill their potential.

CALL FOR EXPERTS

In the Call for Experts scenario, an organization has similarly concentrated behavioral expertise, but there are also prompts and resources that allow this expertise to be integrated more into 'business as usual'. At its simplest, this might mean that standard procedures prompt staff to recognize that the expertise may be needed (e.g., any new requirement in an application process needs to be assessed for its likely effects on behavior). Expertise is not widespread, but access to it is.

At another level, the organization may have invested to ensure that the ability to randomize has been built into new and existing delivery systems, thereby allowing the team to run experiments when they are called on. If working well, this setup would mean that processes stimulate demand for behavioral expertise that the central team can fulfill. That team may also have the institutional support to proactively monitor activities and respond quickly to specific crises.

One benefit to this kind of setup is that it allows teams to select the most promising collaborations, rather than taking whatever is on offer. For example, the team in Employment and Social Development Canada's Innovation Lab claims that a 'careful selection process is critical to the success of incorporating behavioral insights into an organization', since it identifies partners who are open and willing to innovate, which makes it more likely that the subsequent project demonstrates the true value behavioral science can add.[172]

BEHAVIORAL ENTREPRENEURS

In this situation, there is behavioral science capacity distributed throughout the organization, either through direct capacity building or recruitment. The distribution of expertise can work if there are effective support networks and efforts at coordination.[173]

The problem with the behavioral entrepreneurs scenario is that organizational processes do not support these individual pockets of knowledge. Therefore those with expertise find it hard to apply ideas in practice, evaluate their effects, share findings, and build learning. For example, reducing "sludge" often 'requires coordination among a number of teams in the organization', which is a problem when teams work in silos and 'it is not acceptable in the organization's culture to interfere with other teams' affairs.'[174]

These are not just hypotheticals. A review of the Dutch policy landscape found that 'most behavioral policy practices have not been deeply institutionalized', and their advancement 'depends on the ambition of individual enthusiasts' (or, as we've called them, "behavioral entrepreneurs").[175] While they can achieve some successes, their lack of institutional grounding can mean that they become jaded and start looking for other options instead.

BEHAVIORALLY-ENABLED ORGANIZATION

We see a behaviorally-enabled organization as one where there is knowledge of behavioral science diffused throughout the organization, which also has processes that reflect this knowledge and support its deployment.[176] This is the most resilient setup, since staff will be applying behavioral science in a deliberate way as part of "business as usual", rather than through special projects.

A behaviorally-enabled organization would bring together some of our previous proposals. It would embed the behavioral lens mentioned earlier into its core functions, including strategy and operations. For example, behavioral science could be integrated into cross-cutting frameworks like the WHO's 'Health In All Policies' to prompt inter-departmental working.[177] It would address the need to "see the system", recognizing that sustainable outcomes are difficult to achieve through isolated changes. Such an organization would be self-reflective, and carefully explore the varying perspectives and experiences of its staff and service users.

While this setup has the greatest opportunity for scale and sustainability, it also requires the greatest investment. We conclude by talking about what kinds of investments are needed.

CHOICES AND PRIORITIES FOR THE BUILD

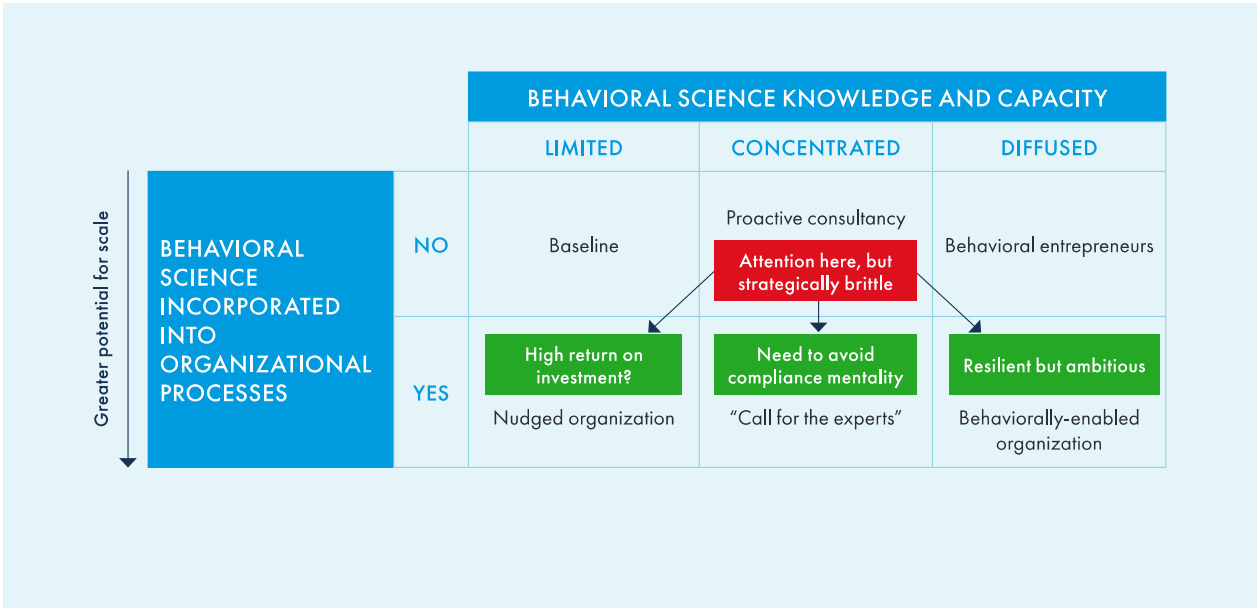
Most discussions make it seem like the meaningful choice is between the different columns in our framework - how to organize your dedicated behavioral science resources. We argue that the more important move is from the top row to the bottom row: moving from projects to processes, from commissions to culture.

A useful way of thinking about this task is about building or upgrading the "choice infrastructure" of the organization, defined as 'the institutional conditions and mechanics of systems - the structures, processes, and capabilities - that directly underlay and support behavioral interventions to help choice architecture solutions work effectively and as planned'.[178]

In other words, we should place greater focus on the institutional conditions and connections that support the direct and indirect ways that behavioral science can infuse organizations. As the diagram below highlights, there are choices to be made about how this is done, based on ambitions and resources.[179]

Working out how best to build the choice infrastructure in organizations should be a major priority for behavioral science. As with many systems, the best option may be to focus on creating the conditions for desired behaviors to emerge, rather than over specifying solutions.[180] But already we can see some features will be crucial.
Dilip Soman and Katherine Yeung argue for the importance of reducing the costs of experimentation, including cheaper data collection, creating an experimental mindset, reducing institutional impatience, and building agility so that organizations can easily adapt to learning.[181] Others have promoted the importance of sharing learning itself, pointing towards the crucial role of Singapore's Civil Service College in 'curating and facilitating an ecosystem of learning opportunities' for behavioral science.[182]

To this list, we want to add new and better ways of using behavioral science knowledge to analyze the behavioral effects of processes, rules, incentives, metrics, and guidelines. Such work has surged recently under the labels of 'behavioral public administration' and 'behavioral operations management', building on a longer tradition of organizational behavior research.[183] We need to ensure that this agenda produces work that has practical value (and not just for the public sector), as in the proposal of "sludge audits" to reduce administrative burdens.[184] Doing so will mean that the behavioral lens we just proposed can be used by an organization's members - and, ideally, by its leaders, who are in a position to achieve broader, systemic change.



03 SEE THE SYSTEM

Many big policy challenges emerge from complex adaptive systems, which present major challenges to the dominant way that behavioral science has been applied. However, we can adapt behavioral science to deal with complexity better, and use it to: identify "leverage points" where a specific shift in behavior will produce wider system effects; understand the collective implications of individuals using simple heuristics to navigate a system; and change the rules of that system to make it more likely that desired behaviors will emerge.

Of course, not every problem will involve a complex adaptive system. So behavioral scientists should first develop the skills to recognize the type of system that they are facing ("see the system"), and then choose their approach accordingly. Fulfilling the broader promise of behavioral science also requires us to expand the ways we tackle problems. The process of identifying targets, exploring drivers, developing solutions, and testing them is strong. The problem is that it contains several assumptions that do not hold when confronting some of the biggest challenges societies face. That's because these challenges often consist of behaviors in complex adaptive systems (CAS).[185]

The risk of talking about 'complexity' is that it may seem like just another way of saying problems are difficult (indeed some use the term as an excuse to do nothing).[186] In fact, we mean applying a particular way of analyzing the world. Our proposal is that combining behavioral science with complexity thinking offers new, credible, practical ways of doing things differently.

First we need to explain complex adaptive systems briefly:

"A complex adaptive system is a dynamic network of many agents who each act according to individual strategies or routines and have many connections with each other. They are constantly both acting and reacting to what others are doing, while also adapting to the environment they find themselves in. Because actors are so interrelated, changes are not linear or straightforward: Small changes can cascade into big consequences; equally, major efforts can produce little apparent change. An important point is that coherent behavior can emerge from these interactions— the system as a whole can produce something more than the sum of its parts."[187]

CAS consist of many different causes, actors, and goals. There are many examples of them in human societies, including cities, markets, criminal justice systems, and political movements. They often create what have been called 'wicked problems', which are ''difficult or impossible to solve because of incomplete, contradictory, and changing requirements that are often difficult to recognize".[188]

Such problems even produce little agreement among different groups about how 'success' can be defined.

The need to understand CAS is becoming more urgent. Complexity has emerged from the sheer number of contacts created by global population growth. At the same time, technological changes like the growth of social media mean that information can be transmitted much faster and cheaply, in an unchanged form, over many different geographies and networks. The fact we do not fully understand how these changes affect human behavior is 'a principal challenge to scientific progress, democracy, and actions to address global crises'.[189]

Indeed, the Covid-19 pandemic, perhaps the most acute global policy challenge of recent times, has several features of a wicked problem:[190]

CONTESTED IDEAS OF SUCCESS

Throughout the pandemic, there has been disagreement between individuals, organizations, and countries about what the overall policy goals should be (suppression, elimination, buying time until vaccines, removing restrictions), let alone consensus about how to achieve those varying goals.[191]

NON-LINEARITY

The initial coronavirus variant appears to have been 'over-dispersed' - outbreaks were seeded by a handful of super-spreading events.[192] It's estimated that around 80% of infections were caused by just 10% of individuals.[193] These properties can explain why some nascent outbreaks fizzle and others take off: viral spread is a non-linear process.[194]

Understanding these features could target policy responses effectively. They might indicate that stopping many repeated contacts within a small set of individuals has little effect on spread, in contrast to reducing random contacts at events and restaurants.[195] In other words, the way the virus interacts with social systems means preventing many social contacts is unlikely to reduce viral spread in a linear way, whereas preventing a few super-spreading events may have an outsized impact.

UNINTENDED CONSEQUENCES

The wide-ranging, intensive actions taken to mitigate the spread of Covid-19 have had many indirect effects. These may have included increases in domestic violence; shortages of hydroxychloroquine in West Africa; reductions in carbon emissions; and changes in healthcare access with shifts to telehealth.[196] Our point is not that these actions should not have been taken per se, but that solutions may change the nature of a problem or create many new ones.

These ideas challenge the assumptions underlying the dominant behavioral science approach

The issue is that 'there are fewer examples of behavioral insights applied to understand behavior in complex change processes'.[197]

Why? Because the realities of complex adaptive systems challenge the main assumptions underlying the dominant behavioral science approach: tight focus on a target behavior, linear effects, and stability.[198] We outline each of these before offering a way forward.

TIGHT FOCUS ON A TARGET BEHAVIOR

People can agree on a specific, measurable target behavior. Interventions that shift this behavior are successful - wider effects are minor and may not be considered, unless pre-specified.

Some behavioral science organizations focus on '"breaking problems down into their constituent parts to understand the desired behaviours'".[199] For some issues, there is value in doing this to identify a core target behavior to drive improvement. But you cannot understand a complex system by breaking it down into parts and mapping it - the way its connections function is key.[200]

We cannot assume that changing a particular part of the system will have the desired overall outcome. For a start, there may be intense disagreement between parties about how a behavior contributes to an issue: some people may see pre-school provision as key to regenerating an urban area; others may view that as an unimportant contributor, compared with reducing crime.

Moreover, focusing on a single behavior to achieve a specific outcome may disguise unintended consequences that hinder progress towards the larger goal. A tight focus on target behaviors can mean seeing only a slice of the picture, and ignoring how actors in a system may respond by producing effects felt elsewhere.[201]

There are many such examples. In the US, grocery stores participating in a government voucher program successfully reduced fraud, but this effort also led to many stores no longer stocking high nutrient foods as a result - ultimately harming recipients' diets.[202] A ban on plastic carryout bags led to 40 million fewer pounds of plastic being used for this purpose, but 12 million more pounds purchased as large trash bags instead.[203] In Brazil, reminders for upcoming credit card payments reduced late payment fees by 14%, but also increased overdraft fees in bank accounts by 9%.[204] Increasing efficiencies in hospitals produces situations where 'activities in one area of the hospital become critically dependent on seemingly insignificant events in seemingly distant areas'.[205]

LINEAR EFFECTS

The intervention affects participants in a direct and linear way, according to a pre-developed theory of change.[206]

In a CAS, actors adapt to the behavior of others, so there is often not a simple relationship between inputs and outputs. Actors in a system may adapt to 'buffer' the effect of an attempted change and keep things apparently stable - i.e., making it seem that the intervention had no effects.[207]

However, repeated efforts may weaken these stabilizing factors, and then a minor subsequent event produces a 'tipping point', where change happens suddenly and the system flips into a new state.[208] An example might be repeated challenges that weaken the commitment of a country's armed forces to democracy, which do not translate into action but which create the conditions for an apparently minor event to trigger a coup.

The point here is that if you do not see your expected behavior in a certain timeframe, in line with a linear 'theory of change', you may assume that it has failed. But, in fact, change may be happening through routes and over timescales you had not anticipated.

STABILITY

You can measure the pre-specified target behavior between point A and point B. The system will remain stable over that time, and people will not adapt in response to the intervention.

Since actors adapt to new conditions, and influence each other in doing so, the nature of the problem may be changed by the introduction of an apparent solution itself.[209] A snapshot of behaviors at one point in time is not enough to claim victory. Perhaps the best example is regulation: market players experiment with and gradually adapt to a regulatory regime, working out how to evade its provisions, until a new policy is needed.[210] Therefore success may actually lie in how well behavioral scientists adapt to the unanticipated effects their own actions produce.[211]

What's particularly sobering for behavioral scientists is that their own field provides reasons why they may be holding onto these principles, despite their limitations. One is the 'reductive tendency', which is 'a process through which individuals simplify complex systems into cognitively manageable representations... when faced with complex concepts, individuals are often inclined to treat dynamic concepts as static, or to generalize across dissimilar domains'.[212]

In other words, behavioral scientists use heuristics based on linear models of change when confronted with complexity, since doing so is 'predictable, comforting, and less mentally taxing'.[213] This tendency can exacerbate another one - an illusion of control. Those who are creating and implementing interventions may overestimate how much control they have over events and outcomes, since they are using an inaccurate mental model of how things work.[214]

The end result, some argue, is a failure to understand how actors are acting and reacting in a complex system that leads policymakers to conclude they are being 'irrational' - and then actually disrupt the system in misguided attempts to correct perceived biases or inefficiencies.[215]

DEVELOP BEHAVIORAL SCIENCE THAT CAN TACKLE PROBLEMS IN COMPLEX SYSTEMS

We have outlined the criticisms. Now, we want to offer hope and a way forward. There is an opportunity to develop behavioral science so it can tackle the aspects of complexity that are common to major policy issues. The starting point is to show how behavioral science can shed new light on well-known features of CAS: the fact that small changes can have big impacts, and the way that actors often use a simple set of rules to navigate a system.

The idea that 'small changes can have a big impact' is often used in the sense that apparently minor features of the way a choice is designed or presented can have a surprisingly large effect on subsequent behavior. Used this way, the idea fits neatly with the standard approach of specific, isolated changes being applied to change a pre-defined behavior.

But CASs show that the statement is true in a different way. They show that 'higher-level' features of a system can actually emerge from the 'lower-level' interactions of actors participating in the system.[216] When they become the governing features of the system, they then shape the 'lower-level' behavior until some other aspect emerges, and the fluctuations continue.

Let's make things tangible with examples. In the Covid-19 pandemic, people were trying to achieve their goals (live their lives) within a broad set of rules and in response to changing events. These adaptive behaviors in specific contexts interacted with the adaptive abilities of the virus. New variants emerged as a result.[217] Some of these variants quickly became widespread and, in doing so, changed the nature of the whole pandemic, most obviously through their greater transmissibility or resistance to vaccines. Policymakers attempting to handle the changed situation then had to come up with new overall strategies.

Experiments have also shown how initial, random fluctuations can emerge and solidify into stark divides that shape societies.

For example, a recent US study showed how partisan policy divisions may actually be produced by these random fluctuations.[218]
The experiment created online 'worlds' where self-identified Democrats or Republicans were asked whether they agreed with up to 20 statements. These statements concerned public issues, but had been selected so they did not reflect pre-existing partisan positions - e.g., whether there should be a move to professional full-time jurors.

In eight of the worlds, participants could see if mainly Democrats or Republicans were agreeing with the proposal. When this happened, strong partisan alignment quickly followed. But the important point is that which proposals fell into the Democratic or Republican camp varied greatly between worlds - sometimes the juror proposal was adopted by one side, sometimes the other. There was initial fluctuation in the initial stages, driven by chance and context, before a sudden non-linear alignment one way or another.

More fundamentally, we can see that norms, rules, practices, and culture itself can emerge from aggregated social interactions. These features then shape cognition and behavioral patterns in turn.[219] The implications here disrupt the crude distinctions of 'upstream' versus 'downstream' or 'high-level' versus 'low-level' policies - or, as one recent paper put it, the "individual frame" and the "system frame".[220] Instead we have 'cross-scale behaviors',[221] where behaviors embedded in specific contexts, shaped by the overall way the system functions, can self-organize and emerge to shape the system itself.[222]

This way of thinking opens up new possibilities. Behavioral science could be used to identify 'leverage points' where behavior could be nudged in a way that produces wider system effects.[223]

Perhaps these interventions could be targeted at the stage of random fluctuations, when contingent features of the context can determine which behaviors get locked in. At that stage, a small, well-timed change could shift events onto a different path. For example, a recent study shows that presenting people with a random selection of opinions (a "random dynamical nudge") from others could prevent segregated echo chambers from forming in online environments.[224]