

## 1 Overview

A growing body of evidence demonstrates that behavioral science insights – research findings from fields such as behavioral economics and psychology about how people make decisions and act on them – can be used to design government policies to better serve the American people.

Where Federal policies have been designed to reflect behavioral science insights, they have substantially improved outcomes for the individuals, families, communities, and businesses those policies serve. For example, automatic enrollment and automatic escalation in retirement savings plans have made it easier to save for the future, and have helped Americans accumulate billions of dollars in additional retirement savings. Similarly, streamlining the application process for Federal financial aid has made college more financially accessible for millions of students.

To more fully realize the benefits of behavioral insights and deliver better results at a lower cost for the American people, the Federal Government should design its policies and programs to reflect our best understanding of how people engage with, participate in, use, and respond to those policies and programs. By improving the effectiveness and efficiency of Government, behavioral science insights can support a range of national priorities, including helping workers to find better jobs; enabling Americans to lead longer, healthier lives; improving access to educational opportunities and support for success in school; and accelerating the transition to a low-carbon economy.

—Executive Order No. 13,707: *Using Behavioral Science Insights to Better Serve the American People*. Signed by President Barack H. Obama on September 15, 2015

In recent decades, psychology and public policy have begun to merge in an unprecedented way. Researchers have worked hand in hand with policy practitioners to improve the design, implementation, and evaluation of public policy. The opening excerpt acknowledges the evidence provided by the field of psychology and sets up a directive to federal agencies in the United States to consider the incorporation of such behavioral insights into their work. It specifically mentions the need for the federal government to design programs and policies that reflect our best understanding of how people engage and make decisions – and social psychology has the disciplinary expertise, influence, and imperative to improve that understanding. While this institutionalization of behavioral insights in government does not represent the first effort of its type, it was certainly one of the most high-profile and wide-ranging in the United States, and similar efforts have been incorporated into government and public institution initiatives globally.

Critically, in this relatively young field, policy interventions and behavioral insights have relied on theory and findings that stem from social psychology. As such, they are commonly built on laboratory-based psychological research, with effects that are often subtle and unstable in the messy and multifaceted real

world of public policy. There is substantial opportunity and need to expand and refine the way that psychological research is produced in order to better enable research and discipline to be more equitably leveraged in applied research and practitioner domains and enable a more inclusive science. Indeed, as one example, the American Psychological Association (APA) guidelines explicitly call on researchers to be “aware of the critical role of science in informing practice and policy and therefore strive to conduct and disseminate research that promotes the well-being of racial and ethnic minorities” (APA, 2019a, p. 27). The APA offers similar guidelines for considering the impact of psychological research on other underrepresented or marginalized communities. Ultimately, this recalibration will allow for psychological science to authentically work in the service of the public good – a goal shared by many, if not most, scholars of the discipline.

In Section 2 of this Element, we offer a definition and explore the history of behavioral insights, with an emphasis on nudging: A nudge is any aspect of a decision context that has an impact on people’s behavior, without removing any decision paths. Behavioral insights, and nudges in particular, have been celebrated as cheap and easy tools to improve efficiency and the cost-effectiveness of outcomes. In the sections that follow, we provide the history and context for several of these successful applications of behavioral insights.

Subsequently, we engage the shortfalls of behavioral insights and contend that the implementation of these insights is built on psychological research that often does not account for, or de-emphasizes, individual difference and distributional effects stemming from personality traits, race, class, and other socio-demographic and cultural factors. In the diverse domains of public policy, it is critical to explore the nuances of these behavioral effects – how and when they operate – in order to engage equity, in addition to efficiency and cost-effectiveness. When we consider the use of these behavioral insight tools by governments and other entities that advocate for and impact a large and diverse set of constituents, the general assumption to date that a small net benefit occurs over a large population is insufficient and can result in inequitable and unethical practices and outcomes. This is particularly true given the broad interest in sustainable and scalable innovations based on insights from psychology. We contend that scholars of social and personality psychology are particularly well-suited to address these gaps in the research and must acknowledge (and account for) several factors in the design of their research.

We propose that the question of whether nudges succeed or fail is far more complicated than what has been examined to date. There are critical equity implications of this limitation. Existing efforts place disproportionate emphasis on whether the desired or hypothesized behavior change occurs

(deemed a success) or does not occur (deemed a failure). For “failed” nudges, researchers consider either no behavioral change or behavioral reactance as failures and generally refocus on adjusting and reapplying psychological tools to change behavior. We propose that psychologists have not focused deeply enough on attempts to integrate the nuanced richness of social psychology into efforts to apply the insights from this field. This is true for scholars across the spectrum of social psychology – those who examine individual differences in decision-making and social cognition and those engaging intergroup, cultural, and social meaning-making processes. Nudges have the potential to carry with them additional psychological costs, and these costs are not always distributed equally. Unlocking the relatively untapped potential of social psychology can help scholars and practitioners begin to unpack and understand these costs.

We articulate some of these consequences and their broader implications for examining and developing equitable research practices and applications. First, we provide an overview and evidence for the psycho-emotional tax that behavioral interventions have the potential to carry. Here, we contend that in addition to examining behavior change, nudges and behavioral insights applications should, in parallel, examine the meaningful costs of stigma, negative emotion activation, and cognitive depletion. Second, we propose that considering these psycho-emotional taxes enables researchers and practitioners to better understand who is being helped or harmed by particular types of nudges, and thereby avoid approaches that provide a disproportionate benefit to those who are relatively better off. Here, we define “better off” as individuals who have the psychological, financial, temporal, or other resources to better reap the benefits of nudges. Third, we call for better data to promote better outcomes for the public good. We underscore the critical importance of disaggregating data and examining behavioral and psychological effects in social psychological research as well as a need to collect data in the wild (i.e., outside of the lab) to enable social psychology to engage with public policy meaningfully and responsibly, with the aim of promoting equitable, in addition to efficient and cost-effective, insights.

Social and personality psychology will be critical for developing a more nuanced understanding of both the behavioral and psychological impacts of behavioral insights approaches and applications. These implications will come alongside a process of developing insights that are better suited to a diverse array of social contexts to which social psychology is being applied. This type of approach to designing and implementing research – one that emphasizes intentionally examining behavioral and psychological outcomes across heterogeneous and diverse samples and contexts – will have significant positive implications for developing equitable behavioral interventions in psychology

and across applied domains. This Element provides concrete recommendations for how the field can contribute in this space.

## 2 Background: A Brief History of Nudging and Applied Behavioral Insights

### 2.1 “Nudge” Defined

First, we review and discuss a brief history of nudging and the evolution of the field of applied behavioral insights. In the literal sense, the definition of “nudge” is to gently touch or push an item or a person. A nudge can be employed to gain a person’s attention or direct their attention in a particular direction. In the world of behavioral insights, the use of the term “nudge” closely follows this literal definition. Richard Thaler and Cass Sunstein popularized the term and use the following definition at the start of their pioneering work, *Nudge*:

A nudge, as we will use the term, is any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting fruit at eye level counts as a nudge. Banning junk food does not. (Thaler & Sunstein, 2009, p. 6)

Critical in this definition is that the intervention must be subtle. In other words, a nudge as opposed to a shove. While this definition has been widely adopted in both academic circles and the popular press, others have attempted to refine the idea. For example, one paper argues that the true definition of a nudge has become a bit confused and attempts to clarify the definition of a nudge (thereby hoping to increase the value of the idea) by revisiting the corresponding foundations in the field of behavioral economics (Hansen, 2016). Hansen more specifically defines a nudge as follows:

A nudge is a function of (1) any attempt at influencing people’s judgment, choice or behavior in a predictable way (1) that is motivated because of cognitive boundaries, biases, routines, and habits in individual and social decision-making posing barriers for people to perform rationally in their own self-declared interests, and which (2) works by making use of those boundaries, biases, routines, and habits as integral parts of such attempts.

(Hansen, 2016, p. 158)

Hansen argues that this definition allows for a more foundational understanding of what it means to engage in this type of influence on behavior. Thaler and Sunstein’s original definition essentially serves as a description of the outcomes that result from Hansen’s. Nudges may operate separately from regulation and

mandates, but they need not be required to do so. It is crucial to note, however, that this definition is not the only recent reconceptualization that has been put forth. One earlier critique argues that many nudges are not truly paternalistic but rather demonstrate persuasion that is arguably rational (Hausman & Welch, 2010). A second posits that nudging is an example of a specific form of governance that can be employed by a policymaker (as opposed to the other four that exist: “hierarchy,” “markets,” “networks,” and “persuasion”). It argues that, typically, nudging does not create sustained behavior change – this would more likely be achieved by promoting social identity and norm changes (Mols et al., 2015). Ultimately, these authors all stress that scholars and practitioners of public policy ought to more diligently consider individuals as complex members of social groups rather than individual actors with cognitive limitations. These two thoughtful approaches are important to note because they underscore the value in considering the philosophical and practical nuance in alternative conceptualizations of this idea. We also acknowledge that it has been disputed as to whether “nudge theory” represents a new approach to behavior change or a reframing of prior efforts to do so. As mentioned, there are worthy arguments about the proper definition (and utility) of nudging, and these perspectives carry implications for both research and practical application of these approaches to social problems.

For our purposes, we apply a straightforward definition of nudging that we believe captures the initial spirit of Thaler and Sunstein’s definition, incorporates some of the subtle nuance offered by Hansen and Mols, and allows for an interpretation that has value to scholars and practitioners alike. It is not our intent to enter into the debate regarding the most appropriate definition of nudges. We consider a nudge to be an element of a choice environment that has the potential to alter an individual’s behavior (whether or not it is intended to do so). This environmental feature must not change the options or decision paths available to the decision maker and must not change the incentives present (economic or otherwise).

## 2.2 Nudging Rises to Prominence

In the early 2000s, several lines of research and other notable events came together to bring the field of psychology to the forefront of popular discussion in an unprecedented way. In 2001, a seminal paper demonstrated that a shift from an opt-in to an opt-out process significantly impacted retirement savings (Madrian & Shea, 2001). This paper performed an analysis of automatic enrollment into 401(k) savings plans. Two crucial findings emerged from this work: first, participation in a 401(k) program was higher when workers were

automatically enrolled; second, many participants who were automatically enrolled kept the default for both rate of contribution and allocation of funds in their account. The authors argue that the tendency of participants to stick to the defaults reflects psychological inertia and a belief that the default provides advice on how to manage the retirement investments. These findings illuminate the reality that behavioral factors may weigh heavily on economic decision-making and provide strong implications for how these types of accounts ought to be managed. More broadly, they also connect to the importance of how defaults are designed and what these decisions communicate when deployed in the public sector – with recent research highlighting that defaults are seen as communicating policymaker preferences and recommendations (McKenzie et al., 2006). Ultimately, Madrian and Shea’s work led to the Pension Protection Act, which is discussed in Section 2.3.

The following year, Daniel Kahneman was awarded the 2002 Nobel Prize in Economic Science for his work with Amos Tversky (sadly, Tversky passed away in 1996, prior to this recognition). Together, Kahneman and Tversky provided fundamental challenges to the role that the assumption of rationality had played in modern economic theory. They established evidence on a wide range of cognitive biases that stem from bounded rationality, as originally defined by Herbert Simon (1957). Tversky and Kahneman brought together the fields of psychology and economics in a way not previously accomplished, and they rooted this work in their observations of real-world behavioral puzzles. Specifically, their extensive work provided insights on judgment under uncertainty. Their studies laid the foundation for what was a new field of research, changing the course and influence of scholarship in both economics and psychology. One of their most prominent contributions, Prospect Theory (Kahneman & Tversky, 1979), incorporated psychological insights as a descriptive theory of choice (compared to Expected Utility Theory from economics) and continues to be applied to real-world decision-making in diverse contexts ranging from labor economics to gambling behavior to decisions about residential movement (Camerer, 1998; Clark & Lisowski, 2017). Kahneman and Tversky’s long friendship and evolution of their work are explored in compelling detail in Michael Lewis’ popular press book *The Undoing Project* (Lewis, 2016).

In 2003, two papers that proved foundational to the amplification of applied behavioral insights were published. Johnson and Goldstein (2003) demonstrated the impact of defaults on decision-making in the context of organ donation rates and made a convincing argument for how policymakers ought to approach the framing of important practical choices. Specifically, they argue that every policy action must have a default specified that will be engaged if no active choice is made. Defaults impose costs (be they physical, cognitive, and/or emotional) on

individuals who must change their status and, as such, must be carefully considered and articulated. Furthermore, they note that the “costs” of switching are sometimes imposed on the group that appears to be in the majority. If the majority of individuals in a society favor organ donation (as measured, in this case, by both a national study and results of the experiment), policies that require active consent are placing the switching costs on the largest group of individuals, potentially amplifying those costs. This was the case in the United States. At the time of this study, the majority of individuals in the United States favored organ donation, and at the same time, many states also had an opt-in organ donation policy, creating decision friction for the dominant opinion. This simple study goes far to illustrate the need for the design and implementation of public policies to carefully consider these costs and the role of defaults.

These findings dovetail nicely with arguments made in “Libertarian paternalism” (Thaler & Sunstein, 2003). This influential paper introduced the concept of libertarian paternalism – “an approach that preserves freedom of choice but that authorizes both private and public institutions to steer people in directions that will promote their welfare” (Thaler & Sunstein, 2003, p. 179). Further, they describe the misunderstandings associated with the concept of paternalism and aim to clarify why it should not be considered a “derogatory” term. Critically, the authors point out that (1) there are often not good alternatives to paternalism, as choice architects are sometimes forced to make decisions on behalf of those who will choose (and often do not want to do so randomly), and (2) paternalism does not always have to involve coercion. In addition, they describe tools that can be used to create a good system of choice. They contend that libertarian paternalism is an option that can preserve individual freedom of choice while simultaneously steering people in the direction that promotes their well-being, and it is often described as a form of *soft* paternalism – meaning paternalism that does not restrict freedom of choice. For example, a government that attempts to curb smoking by mandating images and language regarding diseases caused by tobacco is engaging in soft paternalism. An outright ban on tobacco by the government goes beyond this type of intervention, falling more squarely into the category of outright paternalism.

In their book *Nudge* (first published in 2008), Thaler and Sunstein provide an expansion of their arguments on libertarian paternalism. Specifically, they lay out the principles of effective choice architecture. The existence or nonexistence of a default option is one important point of consideration, but other factors such as the number of options presented and the way that information is described are important features. Thaler and Sunstein use the System I (quick, unconscious, intuitive processes) versus the System II

(slow, conscious, deliberative processes) framework and describe how it leads to many predictable behavioral biases. Ultimately, this leads to policy recommendations (their focus is primarily in the areas of retirement savings and health care/health outcomes). More of the scholarly history of the field of behavioral economics is outlined in Thaler's *Misbehaving* (Thaler, 2016) and many of the important fundamental psychological insights are described in Daniel Kahneman's *Thinking, Fast and Slow* (Kahneman, 2011).

Importantly, while the concept of libertarian paternalism as an intervention approach has received substantial support, it has also received significant criticism. These criticisms have been well developed and articulated by numerous psychologists, legal scholars, economists, political scientists, and philosophers over the decades (e.g., Ewert, 2020; Schmidt & Engelen, 2020). For example, scholars have cautioned that there are clear power dynamics that need to be considered with regard to which entities and individuals are in the privileged position to nudge others' behavior and who is in the lower power position of receiving that nudge. This has implications for the types of nudges that are seen as normative and desirable and has the potential to reify social hierarchies and devaluation of lower status identities, thereby impacting well-being across many metrics. It is a consideration that is particularly relevant to government and public policy, where opting out of engaging with the "nudger" is near impossible given the ubiquitous role of the government in daily life. Moreover, when errors in soft paternalism occur and an entity inadvertently guides individuals to a less desirable decision or behavior, who bears those costs and who corrects for those errors? Because these concerns have been substantially engaged over the last decade, we do not do so here. Rather, these considerations are seeds to the novel concerns we raise in Section 4. For now, we consider the celebrated successes of behavioral insights applications, emphasizing their efficiency and cost-effectiveness.

### 2.3 Public Policy Applications of Behavioral Insights

Next, we detail the applications of behavioral insights to public policy contexts, focusing on efforts to promote efficacy and efficiency, as these have been of primary emphasis to date. Given author expertise, we initially emphasize policy applications in the United States but clarify that there is a positive global trend of many national governments and public entities incorporating behavioral insights into their practices, including in Australia, Singapore, the Netherlands, Germany, and many other countries (Angawi & Hasanain, 2018).

As described in Section 2.2, a key early finding in this field was in the area of retirement savings behavior (Madrian & Shea, 2001). This finding ultimately



led to the 2006 Pension Protection Act. Signed into law by President George W. Bush on August 17, 2006, this law made several provisions to protect retirement accounts, but it also made it significantly easier for employers to enroll their employees into 401(k) plans through an “autosave” feature (Beshears et al., 2010). Specifically, the Pension Protection Act encouraged employers to use automatic enrollment, where employees (after given notice) were automatically enrolled in their retirement accounts unless they explicitly chose not to participate. In addition, employers could make contributions to employee accounts (whether or not the employee chose to participate) or as a match. Finally, contribution rates could be automatically increased over time and those contributions could be defaulted into a diversified portfolio of assets. The Pension Protection Act had broad bipartisan support (it passed in the US Senate with a vote of 93–5 and a House of Representatives vote of 279–131). It was designed based on clear evidence and could be implemented in a fully transparent, nondeceptive manner. Other countries have since adopted similar legislation. To illustrate, the Parliament of the United Kingdom passed the Pensions Act of 2008, which stated that workers had to opt out of the pension plan offered by their employer – as opposed to opting in.<sup>1</sup>

After the ideas from *Nudge* spread, Cass Sunstein received an opportunity to put them into practice in the US federal government. The book *Simpler* (Sunstein, 2014) lays out many of the lessons learned after his appointment as administrator for the US Office of Information and Regulatory Affairs (OIRA) during President Barack Obama’s first term. The OIRA was established as a part of the 1980 Paperwork Reduction Act and is housed within the Office of Management and Budget. Its primary function is to oversee the implementation of government-wide policies that pertain to information technology and privacy policy. Sunstein served from 2009 to 2012 and incorporated many insights from psychology and behavioral economics into this work. In *Simpler*, he maintains that the government can and should be streamlined to improve well-being, through solutions such as simplified administrative processes and improved communication of everyday information.

In 2010, the United Kingdom established the Behavioral Insights Team (BIT), which was also the first government-wide “nudge” unit in the world. The BIT was established within the UK cabinet office to explicitly apply insights from behavioral science across the government. Some of their most prominent work included using letters to increase the payment rate of a vehicle excise tax, nudging a higher rate of payment of fines by sending text messages,

---

<sup>1</sup> See the Pensions Act 2008, Statute Law Database, [www.legislation.gov.uk/ukpga/2008/30/contents](http://www.legislation.gov.uk/ukpga/2008/30/contents).

and using a lottery system to increase voter participation (John, 2014). The BIT was privatized and became a social purpose company in 2014; it now has a global reach – working not only with national governments but also with local authorities, nonprofits, and private entities alike.

Around the same time as the establishment of the BIT in the United Kingdom, a US federal agency was launching its own large-scale effort to explore the application of behavioral insights within its own programs. The Behavioral Interventions to Advance Self-Sufficiency (BIAS) project ran from 2010 to 2016. This work was sponsored by the Office of Planning, Research, and Evaluation in the Administration for Children and Families (ACF) at the US Department of Health and Human Services. This was the first effort by a US federal agency to apply and evaluate the use of behavioral insights in public policy design and implementation. The BIAS project applied behavioral science in the context of ACF-funded initiatives (namely, childcare and work support). Through a collaboration with local implementing agencies and a large team of academic scholars, BIAS tested 15 interventions with nearly 100,000 participants across 7 states (Richburg-Hayes et al., 2017). The BIAS findings included interventions that increased the use of quality childcare by low-income working families, increased frequency of both child support payments and requests for child support order modifications, and increased the rate of childcare subsidy renewals. The BIAS interventions used techniques such as novel communication (e.g., postcards) to prime and remind individuals of actions that they needed to take, designing and delivering assistance to ensure proper completion of complex paperwork, and using identity priming and social norms to increase the perceived desirability of specific actions. The ACF launched the BIAS – Next Generation project in 2015 to build on the success of the fifteen initial BIAS trials. This work has expanded the scope of the original endeavor to explore other areas, such as child welfare and head start programs and working with Temporary Assistance for Needy Families agencies on service delivery.

As the work of BIAS was underway, seeds were planted for an even larger effort within the federal government. The Social and Behavioral Sciences Team (SBST) was chaired by the White House Office of Science and Technology Policy (OSTP). The OSTP worked to explore the potential application of behavioral insights in the US government, and SBST was subsequently established by a 2015 executive order issued by President Barack Obama.<sup>2</sup> This is the

---

<sup>2</sup> See *Using Behavioral Science Insights to Better Serve the American People* – Executive Order No. 13,707 (September 15, 2015), Whitehouse.Gov, <https://obamawhitehouse.archives.gov/the-press-office/2015/09/15/executive-order-using-behavioral-science-insights-better-serve-american>.