Good Practice Principles For Ethical Behavioural Science In Public Policy

Public Governance Policy Paper

OECD

OPSI Observatory of Public Sector Innovation

FOREWORD

Behavioural science draws on principles of psychology, cognitive science, and economics to develop robust and relevant data through testing and experimentation to make policies more effective. Behavioural Insights (BI) began its journey in public policy in 2010 and since then, has played an important role in advancing the use of evidence to inform policy decision-making. **Based on empirical data, BI challenges assumptions about human behaviour by demonstrating how individuals' actions and decision making are not always rightfully categorised as "rational" or "irrational", but sometimes better understood simply as *human*.** The findings of BI are quite specific in that they include evidence about how people respond to risks, how they are affected by inertia, how they can be affected whether a program is "opt in" or "opt out," and how seemingly small features of policy or program design can have large effects on outcomes.

Over the last decade, government officials have been prompted to consider whether BI could enhance their work. **Now that many governments are using BI findings in, and implementing BI functions into, various phases of the policy cycle, there is no better time to prompt public officials with questions of how to apply BI responsibly to ensure the safety, protection, and wellbeing of the public they serve.**

This document contributes to OECD work on enhancing ethics in public policy, which has motivated the development of various publications concerned with advancing best practices in the public sector. These publications have been largely inspired by the OECD's Principles for Managing Ethics in the Public Service (2000), the OECD Recommendations on Open Government (2017), on Artificial Intelligence (2019), on Enhancing Access to and Sharing of Data (2019), in the Recommendations on Agile Regulatory Governance to Harness Innovation (2021), and in other action-oriented OECD tools such as the Good Practice Principles for Data Ethics in the Public Sector (2021). This collection provides a snapshot of the many overarching ethical considerations regulating the work of public officials and governments. This guide draws from these principles to identify the key ethical standards unique and relevant to BI practitioners and government officials seeking to apply BI functions in government.

This guide was produced in direct response to the requests of the global community of BI practitioners in government and was co-designed with members of this community. The OECD has been a frontier for behavioural science applied to public policy, both in the development of tools and standards to serve the BI community and as a convener of this community, offering BI practitioners and policy makers the guidance and support to advance BI in the public sector. In 2019, the OECD launched the Tools and Ethics for Applied Behavioural Insights: The BASIC Toolkit – the first of its kind – an open-source framework designed by BI practitioners and policy makers for applying BI from start to finish.

In continuation of this work and an ongoing commitment to serving the needs of the global community of BI practitioners in government, **this Good Practice Principles guide aims to provide the BI community with a standardised list of the ethical considerations required to apply behavioural science responsibly to public policy.**

ACKNOWLEDGEMENTS

EXECUTIVE SUMMARY

The past decade has seen a rapid increase in the number of governments turning to behavioural science to inform decision making across all policy areas. The use of behavioural insights (BI) helps governments identify the systems and processes influencing human behaviour. Its rise in popularity among governments seeking (for example) to promote public safety and health, or to increase economic growth, has led to a simultaneous increase in the number and types of frameworks designed to help policy makers diagnose and apply BI in their work. These frameworks do well at acknowledging how BI can be applied in various ways to serve a variety of policy goals – whether it be through robust literature reviews, behavioural experimentation and testing, or impact evaluation. In some cases, the use of BI has produced beneficial outcomes for large numbers of people.

Despite a robust and growing set of frameworks and toolkits designed to help government officials apply behavioural science to their work, there are less resources available to practitioners concerned with how to apply these functions responsibly. This leaves the BI community vulnerable to inconsistent and ineffective applications of BI in public policy and demands greater efforts to standardise and systematise ethical processes when apply behaviourally-informed approaches in policy making.

This document is designed to help government officials navigate today's complex policy landscape by providing guidance on how to apply BI responsibly and effectively to tackle today's most pressing policy challenges. For this, we offer a guiding document featuring a "three Ps approach": Principles, Prompts and Practices. It first presents users with overarching ethical principles that reinforce the responsible use of BI in public policy, which is supported by prompting questions, actionable practices, and real-life case studies designed to take users from broad principles of ethics to targeted and feasible actions. It is organised into four sections: 1) Scope: Checking the relevance of behavioural science; 2) Design: Integrating ethics into intervention design; 3) Research and evaluation: Gathering behavioural evidence responsibly; and 4) Policy implementation: Preserving ethics when implementing and scaling BI results for public policy.

In addition to the guide, we offer two complementary tools. Firstly, there is the **ethics checklist**, designed for anyone interested in applying BI to public policy; it offers a clear and simple overview of the key guiding principles for the ethical use of BI. Secondly, there is a **list of prompting questions**, designed for BI practitioners, to stimulate specific ethical deliberations. This list of prompting questions is a useful tool to employ across the policy lifecycle as it motivates deeper reflection to question whether BI is used responsibly.

Not all BI applications follow course from problem scoping to policy implementation. For example, practitioners might use existing evidence to inform policy design without going through new data collection and testing. As such, **this guide, and its complementing tools, are designed to be flexible and accommodate various BI methods and activities.** However, we encourage users to become familiar with all of the included principles, as the standards set out in one section are likely applicable to other phases of BI functions.

TABLE OF CONTENTS

OECD | OPSI Observatory of Public Sector Innovation

CHECKLIST
FOR AN ETHICAL USE OF BEHAVIOURAL SCIENCE IN PUBLIC POLICY

Have you addressed this?

No | In Progress | Yes

Scope
Checking the relevance of behavioural science

1.1 Verify that behavioural science is an appropriate approach for your policy goal

1.2 Establish clear criteria for why the target change should improve public welfare

Design
Integrating ethics into intervention design

2.1 Be transparent with purpose, intentions, and objectives of the intervention

2.2 Consider relevant stakeholders in the design of the intervention

2.3 Set up protocols to identify and mitigate ethical risks
     (such as unintended negative side-effects, both in general and to particular groups)

2.4 Preserve fairness, equality, and dignity

Research and Evaluation
Gathering behavioural evidence responsibly

3.1 Anticipate and plan for unintended consequences (such as backfiring effects or welfares losses)

3.2 Pre-register research questions, hypotheses, and methods before observing the outcomes

3.3 Protect data, privacy, confidentiality, and obtain informed consent when necessary

3.4 Publish and share results

Policy Implementation / Preserving ethics when implementing and scaling BI results for public policy

4.1 Ensure ethical continuity when adapting and scaling

4.2 Ensure communication and implementation guidance with partner institution(s)

4.3 Be accountable and accessible to the public

◇ PRINCIPLE   ? PROMPT   ⚑ PRACTICE

PROMPTING QUESTIONS
FOR AN ETHICAL USE OF BEHAVIOURAL SCIENCE IN PUBLIC POLICY

◇◆ **Principle**

1. SCOPE: Is behavioural science an appropriate approach for your policy goal?

1.1 Verify that behavioural science is an appropriate approach for your policy goal

- What is the policy problem you are seeking to solve?
- How would your solution improve the current situation and increase human welfare?
- Does your policy problem have a behavioural component for which BI can help?
- What is the behavioural change you want to achieve and why?

1.2 Establish clear criteria for why the target change should improve public welfare

- Did you establish clear criteria for why the behavioural change has a positive outcome for the affected population?
- Are these criteria monitored and evaluated regularly?

2. DESIGN: How can you ensure an ethical intervention design?

2.1 Be transparent with purpose, intentions, and objectives of the intervention

- Is the intervention and its purpose as transparent as possible given the desired outcome of your project?
- Are you comfortable with every step of the intervention being as publicly observable as possible (insofar as it does not share confidential or restricted information)?

2.2 Consider relevant involvement of stakeholders in the design of the intervention

- Did you engage stakeholders from the affected population to assist on the design of the intervention? If so, was this group diverse enough to include the most relevant, vulnerable, underrepresented, and marginalised groups?

2.3 Set up protocols to identify and mitigate ethical risks

- Is the criteria set for identifying, assessing, and monitoring risks comprehensive enough to confidently avoid and/or manage risks?
- Are you regularly assessing and monitoring risks throughout the project lifecycle?
- Are there sufficient protocols to avoid or reduce those risks?
- Have you established procedures or protocols to initiate if negative consequences arise?
- Do you have the ability to consult with an internal or external ethics review board, when relevant? If so, have you submitted the project for their review?

2.4 Preserve fairness, equality, and dignity

- Does the intervention promote discriminatory or offensive behaviours?
- Does your intervention stigmatise disadvantaged or marginalised populations?
- Did you take precautions to prevent unfairness in your sample population and in randomisation?

3. RESEARCH AND EVALUATION: How to research and generate behavioural evidence for public policy purposes responsibly?

3.1 Anticipate and plan for unintended consequences

- Have you referred to existing literature and experimentation that may help you anticipate and address unintended results or consequences, in advance?
- Have you established procedures or protocols to initiate if negative consequences arise?

3.2 Pre-register research questions, hypotheses, and methods before observing the outcomes

- Have you pre-registered your research, including observational and experimental research?

3.3 Protect data, privacy, confidentiality, and obtain informed consent when necessary

- Have you gathered the necessary consent from participants that preserves their autonomy?
- Did you disclose all information possible to inform participants while preserving the integrity and outcome(s) of the intervention?
- Are there sufficient measures in place that prevent the misuse of confidential information and data?

3.4 Publish and share results

- Are you and your partner institution(s) prepared and willing to share all results, including negative and null results?
- Are you inaccurately interpreting or mis-representing data?
- When possible, do you publish your research, including observational and experimental research?

4. POLICY IMPLEMENTATION: How to preserve ethics when scaling BI results?

4.1 Ensure ethical continuity when adapting and scaling

- Have you considered new ethical concerns resulting from scaling and adapting in new contexts?

4.2 Ensure communication and implementation guidance with partner institution(s)

- Have you provided partner institution(s) with advice and/or guidelines for implementation and scaling?
- When possible, do you actively communicate with individuals through feedback channels?
- Are the feedback channels accessible and managed in a timely manner?

4.3 Be accountable and accessible to the public

INTRODUCTION

The use of behavioural science and the integration of behavioural insights (BI) in public policy processes enables governments to leverage valuable knowledge about human psychology to inform policy and decision making. Understanding the cognitive drivers behind human behaviour is critical for designing effective policy that is beneficial to and accepted by those it seeks to serve. As an innovative policy-making tool, BI use empirical findings about how people actually think and act, and asks policy makers to put forward those findings to help solve the key policy challenges of today.

Despite the many benefits of behavioural science, particularly when applied to public policy, the potential dangers of an irresponsible or unethical use of behavioural science have contributed to concerns and questions about its use, ranging from accusations of excessive paternalism to suspicions of disguised manipulation. Misuse of BI in the private sector has exacerbated these concerns and has subsequently led to the development of research and discussions into the irresponsible use of BI, such as 'sludge' tactics or dark patterns. **As such, the use of BI by governments has sparked discussion on how we integrate and frame ethical considerations in the design and implementation of policies that aim to change human behaviour.**

Although much has been accomplished in providing evidence for why BI should be used in public policy, more focus needs to be placed on how BI should be used, and more specifically, how can BI be used responsibly. While many are already answering this question through their own processes and best practices, greater efforts to standardise and systematise these practices are key in alleviating concerns about the misuse of BI in public policy. While there is no evidence to suggest that governments are misusing BI, **this guide can act as an additional safeguard against irresponsible or unethical applications of behavioural science when used by government.**

This Good Practice Principles guide provides a set of standards for using BI in public policy that helps us decide how we ought to act in a range of situations. In a sense, we can say that ethics is all about making choices, and about providing reasons why we should make these choices. These standards help us navigate the sometimes blurred lines of appropriate and inappropriate means of generating and using data. Even when intentions and outcomes are clearly beneficial to the public, the path towards these outcomes are not always as clear.

Although academic debate has frequently tried to explore the ethics of behavioural science (Costa, E., & Halpern, D. 2019[2]; Barton & Grüne-Yanoff, 2015[3]; Sunstein, 2015; Hansen & Jespersen, 2013[2]; Thaler & Sunstein, 2008[6]), there are few practical offerings available to practitioners and public officials seeking to systematise ethical consideration when applying BI functions into their policy-making processes. To address this the OECD developed the Tools and Ethics for Applied

Behavioural Insights: The BASIC Toolkit, providing the first ever open-source process framework built by BI practitioners and policy makers for applying BI from start to finish to any policy problem, which included an initial set of ethical guidelines for applying BI responsibly (OECD, 2019[7]). Furthering on these ethical considerations, the OECD in collaboration with the French government, has developed shared guidance on the ethical use of behavioural science in public policy, which can help OECD member and partner countries advance towards a more responsible use of behavioural science for the public sector.

The guide offers practical advice on how to translate broad concepts of good principles for the appropriate use of BI into actionable practices that can be initiated throughout the policy cycle. We draw on practitioners' and academics' own reflections and experiences to include concrete examples of good practices from countries around the world. Revisiting these considerations with a narrower focus on ethics demonstrates how establishing grounds for BI approaches means more than decisions for smart policymaking, but decisions for responsible policy making, by prompting ethical deliberation at each phase of the policy-making process.

◇ **PRINCIPLE**   ◇ **PROMPT**   ◇ **PRACTICE**

Oftentimes, adhering to the standards that promote the appropriate use of behavioural science is consistent with good practice principles of policy-making more generally. In these instances, it is easy to see resemblances between the principles presented in this guide and those featured in ethical guides for other domains. For instance, being transparent with the public is a principle outlined within this document as well as in OECD legal instruments such as the OECD Recommendations on Open Government (2018[8]), on Artificial Intelligence (2019[9]), on Enhancing Access to and Sharing of Data (2019[10]), in the Recommendations on Agile Regulatory Governance to Harness Innovation (2021[11]), and in other action-oriented OECD tools such as the Good Practice Principles for Data Ethics in the Public Sector (2021[12]). Although being transparent is not an ethical consideration unique to BI, any guide outlining principles for the ethical use of BI in public policy would be incomplete without mention of this principle. As such, we do not seek to present considerations prompted by BI alone, but rather, considerations that are necessary for applying BI in public policy ethically.

Similarly, some practices that appear to be good practices for applying BI specifically may also be necessary for ensuring the ethical applications of BI but are generally recognised as necessary functional practices. For instance, preventing unfairness in randomisation is important for the integrity and validity of research results and therefore, can be considered a good functional practice. This practice also promotes principles of fairness and equality, and therefore, contains an ethical dimension. Again, this guide does not encourage users to distinguish between the two. Instead, it accepts and even promotes the notion that the functional application of BI is often also an ethical application of BI.

The main audience for this publication consists of BI practitioners and policy makers working in or with government seeking to or already applying behavioural science to policy-making processes. The principles outlined in this guide build on the OECD BASIC Toolkit to dive deeper into the most common ethical considerations faced by BI practitioners working with(in) the public sector (2019[7]). This guide was also largely inspired by the FORGOOD framework (2020[13]), developed by Liam Delaney and Leonhard K. Lades, aimed at prompting policy makers with questions to identify potential ethical problems in policy design. These good practices contribute to the current offerings of advice practitioners follow when looking to initiate their own BI projects.

This guide helps navigate and respond to the most common ethical uncertainties related to using BI in public policy-making from technical practices for avoiding unfairness in experimentation, to protocols for avoiding backfiring or negative consequences caused by behavioural changes. It is meant to equip practitioners with a step-by-step approach to avoid the misuse of BI by drawing on international good practices according to four sections. Each section contains the three Ps: Principles, Prompts and Practices. The three-P-approach aims to take complex ethical concepts and translate them to actionable goals that are supported by real-life examples of how ethical principles are upheld within governments and organisations. The four sections are:

 SCOPE: Checking the relevance of behavioural science;

 DESIGN: Integrating ethics into your intervention design;

 RESEARCH AND EVALUATION: Gathering BI evidence responsibly; and

 POLICY IMPLEMENTATION: Preserving ethics when implementing and scaling BI results for public policy.

This guide is designed to be useful to policy practitioners facing a variety of methodologies, interests, challenges and demands. In recognising that different actors play different roles in the lifecycle of a BI intervention, every section is designed to be considered and applied within the parameters of the project. That is, some attempts to uphold certain principles may be limited due to a lack of capacity or authority to formalise it yourself. Regardless, familiarising yourself with the various ethical considerations associated with BI in public policy, whether under your authority or not, remains useful in any attempt at applying BI ethically.

Those interested in creating a BI intervention from scratch may follow the whole guide from section one to four. Conversely, those conducting exploratory research only, may find the recommendations set out in section three most relevant. Those who are interested in adapting a BI intervention tested in another context may find the last section about the implementation and scaling most applicable to their work. Finally, those who are tasked to provide senior leaders with quick advice applying a BI lens to a given policy issue may focus more on the recommendations presented in section one and four. Regardless of which section serves as most relevant to your project, we encourage you to familiarise yourself with all of the following sections, as well as the ethics principles checklist and the list of prompting questions, as the ethical considerations presented in one section may prove to be of equal importance in additional phases your project's lifecycle.

This publication focuses on ethical considerations to the application of BI in public policy and, in doing that, there are limitations to its focus. It is useful to keep in mind the following:

This guide **does:**

- Offer a (non-exhaustive) set of good practice principles to support the ethical application of BI in public policy

- Act as a guiding tool to prompt ethical deliberation during a BI project which can be used completely or partially, and in various orders

- Provide a starting place to have richer discussions about the ethical considerations of BI experimentation and interventions

This guide **does not:**

- Provide a definitive list that, when followed, automatically assures your intervention is ethical or will yield significant results

- Cover the operational and/or structural conditions required for a successful application of BI in public policy

- Offer any pre-determined solutions to the trade-offs between ethics and efficiency

The application of ethical principles is not binary, and thus, deep consideration must be paid to the layered applications of ethics, paying full attention to the contextual factors that influence how ethics can and should be applied. As such, we recommend that you use this guide accordingly.

1 **SCOPING**
Checking the relevance of behavioural sciences
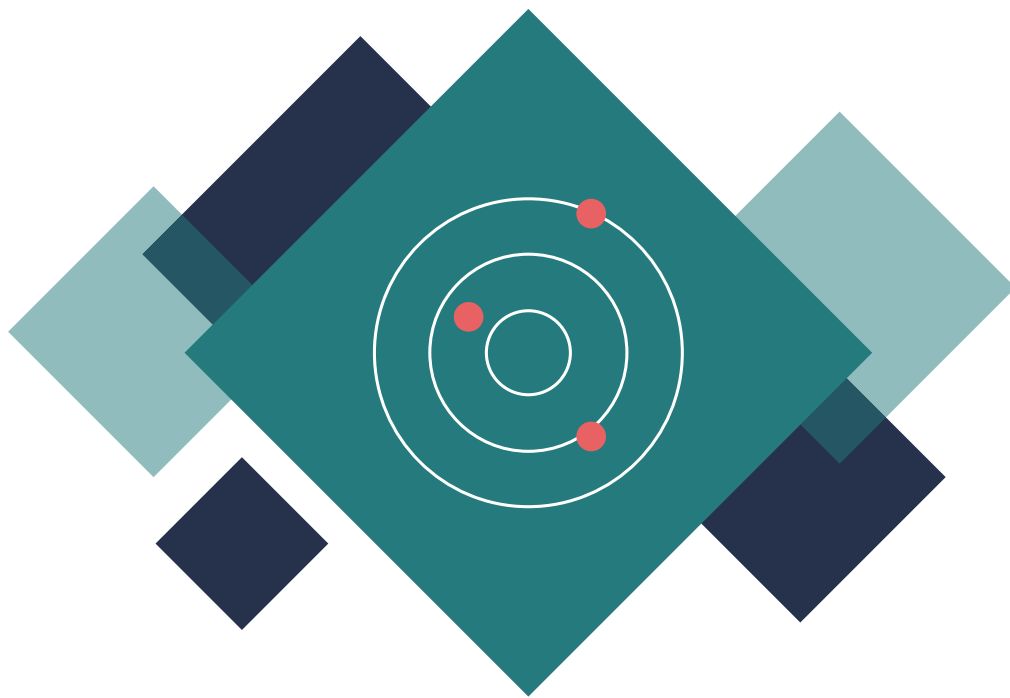
Ethical deliberation should start the moment a BI intervention is proposed. BI approaches rely on insights about individual and group-level behaviours and preferences and leverages that information to inform policy either by encouraging obvious or subtle changes in individuals' behaviours. As such, careful consideration should go into deciding whether encouraging behavioural changes is required for the precise policy goal. Instead of deciding this with intuition, international standards and good practices, such as those outlined throughout this guide, point to the benefits of applying a behavioural lens to problem-framing and solution ideation. Recognizing the limitations of BI in achieving policy outcomes is a key consideration for determining how the use of behavioural science can help advance a policy goal.

This first section targets both BI practitioners and policy makers by offering guidance on how to systematise decision-making processes concerning whether BI is an appropriate policy tool among many, by considering the effect it has on its target population when employed alone or in combination with other policy tool(s).

PRINCIPLE 1.1: VERIFY THAT BEHAVIOURAL INSIGHTS IS AN APPROPRIATE APPROACH FOR THE POLICY GOAL

Using BI to design and implement policies can appear appealing initially due to the alleged cost-effectiveness and its innovative approach (Benartzi et al., 2017[14]; OECD, 2019[7]). However, despite many policy problems containing a behavioural dimension, not all policy problems can benefit from a BI perspective nor should be addressed with BI interventions. But for those that do, behaviourally-informed interventions are key for employing effective policy that is informed and designed according to real human behaviour.

Traditionally, BI is often used to enhance rather than substitute more classic policy-making. For example, utilising BI sich as social norms to complement existing policy approaches such as taxation to increase compliance (Hallsworth et al., 2017[15]; Hallsworth & Kirkman, 2020[16]). This is unsurprising given the complex nature of today's policy problems which demands collaborative and creative solutions that acknowledge human behaviour and decision-making processes. As a policy-making tool that can be employed alone or in combination with others, it is important to identify the appropriate entry points for behavioural applications. Understanding the policy cycle and where behavioural functions fit into larger policy ecosystem can help ensure BI is not inaccurately or misused. As such, when choosing BI as a policy-making approach, you should identify why and where BI fits into the policy process by:

Prompts 1.1:
- **What is the policy problem you are seeking to solve?**
- **How would your solution improve the current situation and increase human welfare?**
- **Does your policy problem have a behavioural component for which BI can help?**
- **What is the behavioural change you want to achieve and why?**

BOX 1.1.1: BETA DISCOVERY TOOL

The Behavioural Economics Team of the Australian Government (BETA) has developed a tool allowing practitioners to better identify and specify the behaviour(s) they aim to target and assess if and how BI can help. It considers a number of factors, such as whether there is previous research to support that the proposed intervention will result in desired behavioural change. It also prompts questions about how the population is informed, whether citizens are able to perform or change the behaviour (i.e., access to infrastructure, competences, resources, etc.), and whether they have beliefs supporting it.

This tool assists its users in better identifying whether this is a behavioural component worth considering and identify the extent to which BI may be useful in obtaining your desired outcome(s).

Source: Behavioural Economics Team of the Australian Government. (n.d.[17]). 'Behaviour Discovery Tool'. Retrieved April 7, 2022, from https://youropinion.au1.qualtrics.com/jfe/form/SV_bJBe80XcOFtpHjD

PRACTICES 1.1:

- **Identifying whether the policy problem has a behavioural dimension.** Clearly identifying the underlying behavioural aspects contributing to a greater policy issue and establishing grounds for why the behavioural science can help target these behaviours is necessary for establishing BI's relevancy in advancing a policy goal.

- **Referring to robust and relevant evidence for the behavioural change.** Exploring what other BI practitioners or researchers have previously discovered on the subject is always useful to understand the key takeaways and good practices that could be used to leverage the behaviour at hand. Ensure criteria is set that guarantees the quality of data and sources used to support your research. It may be easier to advocate for the use of a behavioural method if it has already been proven to yield positive results in previous studies. This may also include collaboration and consultation with other researchers, BI units, or people representative of the target group to further enrich your understanding of the policy issue at hand and to contribute to the robustness of your intervention and policy outcomes.
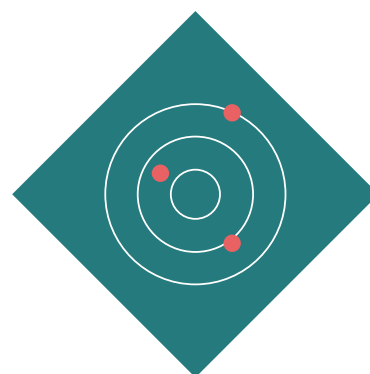
- **Determining the comparative advantage of BI in tackling the policy problem.** In some contexts, policy makers may identify structural interventions as better suited for achieving certain policy outcomes but face economic or political constraints causing them to consider using other tools and approaches (e.g., human-centred design, foresight). In most cases, a careful cost-benefit analysis determining to what extent individual behavioural change versus structural change is required to achieve the policy goal can ensure BI applications are being employed realistically and appropriately. This can be done by using BI toolkits designed to help BI practitioners determine whether BI is the best approach for the policy challenge at hand.

- **Acknowledging how much of the policy problem can be reasonably addressed with BI interventions.** Understanding human behaviour and preferences are vital to creating informed policy. However, they are not the only considerations or factors required to address key policy challenges. Identifying the extent to which BI interventions can advance towards a specific goal over other policy-making instruments is valuable both for advocating its value as a policy-making instrument and for identifying its limitations when dealing with certain key policy areas. In some cases, a false understanding of behavioural science may lead some to believe a structural problem is a behavioural one (or vice versa). Working with key stakeholders to understand what the goal is and how it can be achieved with BI approaches can help reduce instances where BI is poorly or incorrectly applied (cf. section 2.2 below).

BOX 1.1.1: OES GUIDE FOR IMPROVING PUBLIC INTEGRITY

Working with the United States Chief Financial Officers Council, the Office of Evaluation Sciences (OES) at the U.S. General Services Administration developed a Behavioral Insights Guide for Improving Payment Integrity. This guide aims to help government employees assess whether a behavioural approach is a good fit for their program, identify which interventions are appropriate, and implement those interventions with the goal of reducing improper payments, such as payments made by the government to the wrong person or in the wrong amount.

The guide discusses the types of improper payments that are a good fit to address with BI, such as those resulting from unintentional error rather than intentional fraud. It reports the average effect when these types of interventions have been applied by government agencies, and tells readers: "More intensive changes, like passing new legislation or revising the content of forms (as opposed to just making small changes to how the form looks), may have the promise of achieving larger effects, but program staff may not be able to make these sorts of changes themselves…When a small change saves money, and these savings accrue across many payers or payees, behavioural insights interventions can be highly cost effective".

Source: Office of Evaluation Sciences. (2021[18]). A Behavioral Insights Guide for Improving Payment Integrity. United States: US Chief Financial Officers Council. Retrieved April 7, 2022, from https://oes.gsa.gov/collaborations/improper-payments-playbook/

PRINCIPLE 1.2: ESTABLISH CLEAR CRITERIA FOR WHY THE TARGET CHANGE SHOULD IMPROVE PUBLIC WELFARE

Making the case for implementing BI approaches over other policy-making tools alone is not sufficient for securing its success. BI practitioners, together with policy makers, should be concerned with determining how BI can bring positive outcomes for the target population and anyone else possibly affected by the behavioural change. Establishing criteria to monitor the success of the intervention in relation to the policy goals and assessing these throughout the intervention process can help ensure the intervention progresses effectively. The variables chosen to monitor this should be regarded as the most appropriate indicators for monitoring success in the context of the project and target group.

Targeting different actors can have different ethical implications and as such, ethical consideration should be involved in the very identification of the target group itself. For example, a recent OECD study assessing how to use BI to increase the uptake of generic drugs in Chile distinguished between BI measures to increase prescription of generics drugs by physicians/pharmacists as compared to BI measures to increase demand and familiarity of generics among consumers (OECD, 2021[19]). In such instances, it is important to consider the following when determining who to target and why:

Prompts 1.2:
- **Did you establish clear criteria for why the behavioural change has a positive outcome for the affected population? Are these criteria monitored and evaluated regularly?**

PRACTICES 1.2:

- **Identify the specific behaviour(s) that require(s) change.** This can be done using frameworks that help determine the behavioural aspects of a given problem and the diagnostic indicators for identifying the behaviour you seek to change, such as the Behavioural Insight Team's (BIT) MINDSPACE and BIT Barrier Identification Tool, the World Bank's Mind, Society and Behaviour Report , the OECD's BASIC framework, or by using frameworks for specific policy areas such as the OECD's Guide for Behavioural Insights for Better Tax Administration or the Dutch Government's Communication Activation Strategy Instrument (CASI) for applying BI in public communications.

- **Ensure that citizens have the capability and motivation to change their behaviours** otherwise, any attempt to change such behaviours while respecting and preserving individual autonomy will likely be ineffective.

- **Align BI interventions with the public's interests.** Agents of BI must make efforts to consider and incorporate the public's best interests into every phase of the intervention, from ideation to implementation. This requires using recent data and literature to determine the desirable outcomes, identifying public support for leveraging BI approaches for the policy

goal, and evidence to support how the individual behavioural change is conducive to greater common good (cf. Box 1.2).

- **Make sure data is relevant for the target population.** Maximising the relevance of your data, is crucial for BI research to account for the target population when testing. This means collecting up-to-date and reliable data by testing on representative samples of the target population. Randomisation of participants can only partially account for this, since the sample itself, and the data that derives from it, might be biased (e.g., testing an intervention addressing the entire population with a sample of students). Applying an intervention on a population different from the one that has been tested may cause unforeseen effects and infringe on the evidence-based principles of BI (Hallsworth & Kirkman, 2020[16]). BI practitioners may not always be able to test on representative samples, but their recommendations for policy makers should include the limitations of BI to avoid potential harms to the target population and negative spillover effects to others during implementation. Additional efforts should be made to include the most relevant, vulnerable, underrepresented, and marginalised populations.

- **Use detailed and predictive modelling** as a useful tool for anticipating how different groups of citizens will react to the introduction of a new measure, and for comparing the impact of alternative policy measures. A related example can be found in a 2019 OECD study assessing incentives to reduce traffic congestion in Israel, which included an analysis of how different congestion charging systems would impact the behaviour of different categories of citizens (OECD, 2019[26]).

- **Work with necessary stakeholders to plan for implementation and scaling early in your project's timeline.** Effective preparation for scaling results may help avoid irreversible ethical errors later on in the timeline (cf. boxes 4.1.1, 4.1.2 and 4.1.3). This can be achieved by coordinating with the relevant actors in establishing the goals of the experimentation, including intentions for implementing and scaling interventions, and in the ideation phase by identifying potential ethical threats and addressing these when designing the intervention. This can ultimately reduce uncertainty and risk while enhancing ethical considerations, resource allocation and preparedness of foreseen barriers or challenges (see Section 4 for more).

- **Determine quantifiable variables** on which the behavioural change can be reasonably believed to have a positive impact on the target audience based on research and consultations. These variables should be determined at the earliest stages of the project but can be refined a posteriori. Developing these variables with decision makers, if done properly, establishes indicators for success that can be assessed, monitored, and evaluated throughout the project.

A 2011 systematic review of behavioural change frameworks introduced a method for characterising behaviours and identifying appropriate interventions to stimulate desired behavioural change. The framework identifies three main conditions for the behavioural interventions may be used: Capability, opportunity, and motivation. These three conditions form the central component of the "behaviour change wheel" which is designed to help link sources of behaviours and interventions that causes changes in those behaviours (Michie et al., 2021[25]).

Inspired in part by this, the OECD created the ABCD Framework as an additional tool of the OECD's BASIC toolkit. The ABCD framework is designed to help identify and analyse relevant behavioural blockers and barriers. It is divided into 4 sections that start with broad areas of potential behavioural biases and bottlenecks (ABCD: Attention, Belief formation, Choice, Determination) to identify specific examples of real behavioural interventions which can be implemented to address these biases (OECD, 2019[7]).

BOX 1.2.1: UNETHICAL VERSUS ETHICAL USES OF BEHAVIOURAL INSIGHTS

One aspect for upholding the ethical use of BI for public policy explored in Section 1 is ensuring the outcome(s) of the behavioural intervention support the public good. Any application of BI that is not primarily concerned with increasing the wellbeing of the public immediately calls into question the rightful use of BI for a given outcome. As government officials tasked with designing policy to improve public welfare, the public must be the primary beneficiary of all policy decisions. However simple this notion is, decision and policymaking processes are less clear about how to determine and monitor public wellbeing and how exactly to decipher who should be targeted by a specific policy and when.

There are clear cases in which behavioural science can be used in opposition to the wellbeing of its target audience. For instance, sludge tactics are often employed to discourage individuals from making choices that are in their best interest, usually to benefit others (Sunstein, 2018[27], 2020[28]; Thaler, 2018[29]). Consider a subscription service that allows you to subscribe to their services simply by clicking a box online. The service may choose to impose additional barriers like having to confirm your identity, log into your email, or type a reason why you want to unsubscribe to a service to dissuade users from completing a task by adding additional frictions. This may be intentional or unintentional, but consequently encourages the suboptimal choice for the user when choosing to unsubscribe.

Something similar can be said about the use of dark patterns, which exploit individuals' biases to encourage decisions that are not in line with individuals' best interests. A common example of this is using BI to predict when consumers are most likely to make an online purchase and designing marketing and advertising to target this user and increase their chances of making unnecessary purchases (OECD, 2020[30]; OECD, 2021[31]).

Both sludge and dark patterns serve as clear examples whereby the application of BI would not be conducive to the public's best interests and therefore, cannot be considered ethical practices. However, even policies with good intentions can be unethical. As such, determining what is required to enhance public welfare based on empirical evidence and implementing policy according to rigorous testing in an open and accessible way can help alleviate concerns of the unethical use of behavioural science in public policy (OECD, 2017[32], 2019[33]).

2

DESIGN
Integrating ethics into intervention design
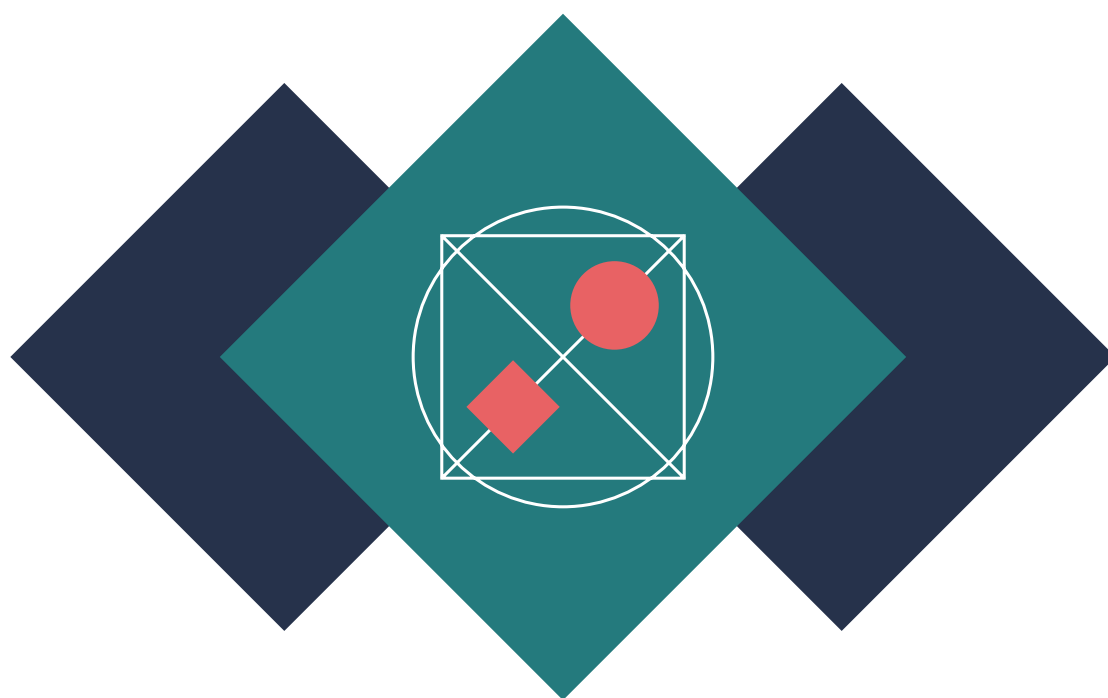
In the previous section, readers are asked to consider the various conditions under which adopting a behavioural approach would be appropriate. The following section presents the ethical principles worth considering once you decide to proceed with behavioural experimentation. Although the intervention has yet to be initiated, the design phases of an intervention present practitioners with the opportunity to engage in important discussions with partner institutions, decision and policy makers, advisors, and other key stakeholders to establish the objectives of the intervention, including the policy implications of the anticipated results, and the possibility for scaling or implementing the intervention.

PRINCIPLE 2.1: BE TRANSPARENT WITH PURPOSE, INTENTIONS, AND OBJECTIVES

Transparency as a guiding principle of public policy, promotes openness and accountability to the public that is being served through any policy. However, transparency in the realm of behavioural science speaks to how subtle or obvious the intervention is to its target audience. For example, Hansen and Jespersen (2013[5]) use a behavioural intervention matrix to classify different types of behavioural interventions according to the dual process theory – such as System 1 or System 2 thinking introduced by Daniel Kahneman in his book Thinking, fast and slow (2011[34]) – and to what extent the target population can detect the presence and objective of the behavioural intervention, categorised as either 'transparent' or 'non-transparent' (Hansen & Jespersen, 2013[5]; Kahneman, 2011[34]).

Although the distinctions between these two types of transparency are important to keep in mind, behavioural practitioners must be mindful of both, as neither are irrelevant to the work of those applying behavioural science to public policy. Transparency in both instances defends the notion that one should never intentionally deceive or manipulate individuals with malicious or harmful intent and when possible, efforts should be made with the relevant actors to share whatever information possible, notably the purpose, intentions, and objective of the intervention. As such, relevant stakeholders should aim for as much openness as possible to encourage accountability and awareness in the intervention processes.

Prompts 2.1:
- **Is the intervention and its purpose as clear and open as possible given the desired outcome(s) of your project?**
- **Are you comfortable with every step of the intervention being as publicly observable as possible (insofar as it does not share confidential or restricted information)?**

PRACTICES 2.1:

The following should be considered according to what is possible given the parameters of your project:

- **Aim for choices that preserve personal autonomy, rights, and freedoms as much as possible** when designing an intervention. Encouraging policy outcomes that recognise that individuals have differing preferences by allowing them to make their own choices is the best way to preserve individual autonomy and avoid accusations of manipulation or paternalism.

- **Be as open as possible** by disclosing the underlying mechanisms of the intervention, when appropriate and useful to the relevant stakeholders (Bruns, 2021[35]). This extends to openness in research methodologies for scientific rigor and scrutiny, as well as transparency and openness with those either directly or indirectly affected by the intervention's outcome(s).

- When implementing non-transparent or unavoidable interventions, **advocate for citizen feedback channels that are easily accessible.** Avoid instances where citizens are punished or penalised for their choices and strive to clearly communicate as much information about the intervention as possible, to the relevant stakeholders.

- When appropriate, **test both the intervention's efficacy** (i.e., in terms of behavioural change), **and the participants' understanding of the intervention** (e.g., via interviews or feedback questions) to have an accurate perception of the extent to which the intervention was transparent and avoidable (more on this in section 4).
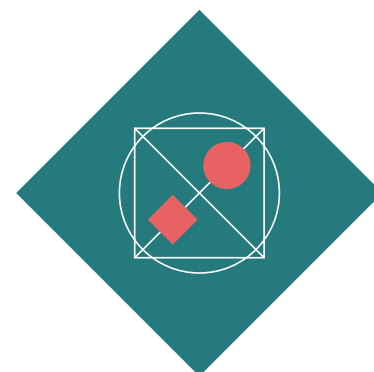
BOX 2.1.1: ORGAN DONATIONS AND AVOIDABILITY

Consider an intervention with the goal of increasing organ donations. There are a variety of ways to increase enrolment in organ donation schemes such as media campaigns (Arizona, United States), simplifying registration processes (Ontario, Canada), or default opt-in schemes (Sweden). Each option has a varying degree of effectiveness as well as varying degrees of avoidability. That is, a brochure prompting individuals to consider enrolling in the scheme is a more avoidable, or perhaps passive, approach to increasing enrolment than a default approach that automatically enrols citizens into the program unless they 'opt-out'.

When comparing various behavioural levers, consider which approach empowers the target group to be informed decision makers and best preserves their individual autonomy. Consider naming this as a desired outcome of the intervention when scaled or implemented, or integrating this into policy recommendations and briefs when possible.

Sources: Bramhall, S. (2011[36]); Ontario Ministry of Health. (2016[37]).; Zion & Zion. (n.d.[38]).

BOX 2.1.2: PROMPTING ETHICS IN BIG AGREEMENTS

British Columbia's Behavioural Insights Group integrates a statement about ethics in all its contracts with the ministries they serve, as well as with their partners. It specifies BC BIG will apply an ethical approach to all stages of their projects and that their recommendations will be based on ethical considerations.

Formalizing these commitments is a useful way for ensuring ethical continuity in all BI projects and contributing to best practices and standards within behavioural sciences.

Source: Devine, H. (2021, August[39]). OECD Interview with Behavioural Insights Experts Working in Government [In-person

PRINCIPLE 2.2: CONSIDER RELEVANT STAKEHOLDERS IN THE DESIGN OF THE INTERVENTION

Since BI practitioners and policy makers cannot anticipate all negative side effects of an intervention, including a representative set of stakeholders – either formally or informally – can enrich your understanding of the intervention, including its effectiveness, unintended outcomes or backfiring effects, or undesired discriminatory effects (Lades & Delaney, 2020[13]). You may choose to engage stakeholders at fixed stages of the project with exercises that allows those involved to brainstorm and identify the potential risks associated with your project, or informally throughout the project's lifecycle or both, if necessary.

Prompts 2.2:
- **Did you engage stakeholders to assist in the design of the intervention?**
- **If so, was this group diverse enough to include the most relevant, vulnerable, underrepresented, and marginalised groups?**

PRACTICES 2.2:

- **Consider those who are most concerned by the target behaviour(s),** including those represented by the target group and, when appropriate include them as early and frequently in the interventions lifecycle as possible to ensure their input is integrated in the scoping phases onward.

- **When appropriate, include vulnerable, underrepresented, or marginalised groups** to ensure representation and diversity in problem framing, solving, and intervention design (OECD, 2017[8]).

- **Grant contributors equal and fair opportunities** to be informed and consulted, and to be actively engaged in all relevant phases of the policy-cycle (OECD, 2017[8]).

Your panel of contributors could include representatives of the target group or experts familiar with the

outcome(s) and effects of the behavioural change. However, you may also wish to include the following actors depending on the parameters and goals of your project:

- Public entities and civil society;

- Those primarily affected by your intervention;

- Experts in BI, psychology, policy making, or other related fields;

- Public officials from other departments or governments;

- Academics familiar with BI, your research, or your project's outcome(s);

- Private sector actors familiar with BI, your research, or your project's outcome(s);

- Other affected parties or experts involved in the policy-making area.

BOX 2.2.1 FRAMEWORK FOR COLLABORATIVE APPROACHES TO POLICY DESIGN

The Open Government Partnership (OGP) offers their OGP Participation and Co-Creation Toolkit outlining international best practices for establishing effective and lasting collaborations between governments and the private and third sectors. The toolkit features the contributions of 39 countries mapping out 100 case studies on partnership and co-creation policy design.

Source: Varga, P. (2018[40]). 'OGP's Participation and Co-Creation Toolkit: From usual suspects to business as usual'. Open Government Partnership. Retrieved April 7, 2022, from https://www.opengovpartnership.org/stories/ogps-participation-and-co-creation-toolkit-from-usual-suspects-to-business-as-usual/

PRINCIPLE 2.3: SET UP PROTOCOLS TO IDENTIFY AND MITIGATE RISKS (SUCH AS UNINTENDED NEGATIVE SIDE-EFFECTS, BOTH IN GENERAL AND TO PARTICULAR GROUPS)

Adapting behavioural approaches within policy processes will likely raise varying values and interests that must be carefully balanced by all stakeholders. BI practitioners have the ability to anticipate and address the complex ethical considerations of BI interventions. Striving for collective action in evaluating the potential ethical risks at each step of the project's lifecycle can ensure that anticipated and unanticipated ethical concerns are not overlooked. For this, BI practitioners should:

Prompts 2.3:
- **Is the criteria set for identifying, assessing, and monitoring risks comprehensive enough to confidently avoid and/or manage risks?**
- **Are you regularly assessing and monitoring risks throughout the project lifecycle?**
- **Are there sufficient protocols to avoid or reduce those risks?**
- **Do you have the ability to consult with an internal or external ethics review board, when relevant? If so, have you submitted the project for their review?**

PRACTICES 2.3:

- **Review relevant literature relating to your research project design.** Exploring what ethical considerations other BI practitioners or researchers encountered during their projects is a valuable contribution to the quality and outcome(s) of your research and intervention design. This may include consultation with other researchers, experts, or practitioners to further enrich your understanding of the ethical considerations at hand and to contribute to the design of your project and impact of your intervention when scaled.

- To the extent possible, **identify the specific ethical aspects of the issue at hand.** Working with stakeholders to identify the ethical considerations of an intervention can help ensure all parties have a shared understanding of the intentions, objectives and limitations of the intervention and its design.

- **Engage stakeholders and partners to establish definitions, criteria, and levels of risk** potentially associated with the project based on previous literature reviews and consultations. This should be informed by any foreseeable harms for the sample population, target population and general public (i.e., physical, psychological, economic, legal, social, or political harms, etc.). Additional criteria should be established if your study involves vulnerable groups, such as minors, cognitively impaired groups, marginalised groups such as refugees or indigenous people.

- **Work with partners to develop and integrate protocols and procedures** to frequently assess the risk factors associated with the project based on reputable and measurable variables for risk assessment and monitoring.

BOX 2.3.1: ETHICS REVIEW BOARDS

Ethics review boards are vital for enforcing and maintaining the best practices and ethical principles that preserve the integrity and credibility of behavioural science. The ethical clearance provided by an independent review board is the best way to ensure to all relevant stakeholders that ethical precautions have been rigorously considered throughout the research process (Lades & Delaney 2020[13]; OECD 2019[26]). Review boards sometimes exist within institutions themselves and practitioners may have to solicit their approval. Even in these cases, it is important to ensure BI-specific competences are included as the guidelines and procedures for usual research in public policy may not be sufficient. In order to ensure a BI project is reviewed by the proper authorities, BI practitioners should, when appropriate:

- Advocate for an independent review board that has **BI-specific competencies** that allow them to understand and consider the ethical considerations unique to behavioural sciences.

- Ensure the ethics review board **endorses all relevant and pre-existing best practices and ethical standards** such as those provided in this guide and those outlined in the American Psychological Association's Code of Conduct (2003[41]).

If BI practitioners are responsible for the ethical assessment and cannot consult an internal research review board, they should:

- Make sure the internal research review board has access to **at least one person with BI experience** and awareness of ethical issues associated with BI.

- **Consult with an external ethics committee** that can provide relevant advice. Ensure they have experience and competence with research in social, psychological, or cognitive sciences.

- Design a method to **assess whether a project requires a specialised ethical review**. Examples of special cases may include, testing interventions which involve unconscious or involuntary mechanisms, recruitment of participants from vulnerable populations (e.g., children under the age of 18), or the experiment involves intentionally withholding the purpose or outcome(s) of the intervention from participants and does not seek consent before the intervention (more on this in Principle 3.3 below).

Even in cases where risks are not perceived as high, it may still be worth consulting with an ethics review board with knowledge of behavioural science as an addition to your stakeholder group. If deciding to undergo either internal or external ethical review, consider engaging them early in the project's lifecycle to integrate their ethical perspective from project initiation.

Source: Lades, L., & Delaney, L. (2020[13]); OECD. (2019[26]).

- **Establish why the risk is justifiable** for the project's goals and outcome(s). If this has already been determined, ensure that the design and progression of the project aligns with its justification and alert the relevant actor(s) if this reasoning becomes threatened by changes in the intervention.

- If proceeding with an intervention with high potential risks of consequences, **work with relevant stakeholders to establish indicators to monitor risks** and determine protocols to initiate if and when necessary. To do this, it is helpful if you clearly identify the following questions:

  » Who (institution and actor) is legally accountable for the intervention's outcomes?;

  » Who has the authority to change or terminate the intervention design if needed?; and

  » Who should be monitoring and reporting on risk factors, how often, and to whom?

- If potential risks are perceived as high, seek external and/or independent ethics review (cf. box 2.3.1).

BOX 2.3.2: RISK ASSESSMENT TOOLS FOR BEHAVIOURAL RESEARCH AND INTERVENTIONS

Each project undertaken by BETA goes through a risk assessment procedure to determine whether further consultation of their partner academic ethical board is required. Their assessment covers the most important risks, such as foreseeable harms of different sorts, involvement of vulnerable groups, and topics specified in the National Statement (i.e., pregnancy, disclosure of information, cultural, political, religious issues, addiction, and sexual identity).

The OES at the U.S. General Services Administration developed the Behavioural Insights Guide for Improving Payment Integrity which includes templates designed to illustrate the considerations involved in applying a behavioural insights approach to reduce improper payments. One of the included templates helps users prioritize the goals and objectives of their intervention while prompting them to consider the associated risks of their design.

To assist BI practitioners, ministry clients, and partners in their commitment to adopt and ethical approach in all stages of their project, BC BIG has developed an Ethics Impact Assessment Tool. This tool does not help users determine whether a project it is ethically sound or not. Instead, it guides practitioners through a series of questions aimed at eliciting reflection around fundamental ethical considerations for BI projects: whether participants' informed consent is necessary, and assessment of the risk to participants, safeguards to ensure privacy, respect, well-being, and autonomy of participants. The tool drew upon the guidelines and principles from the Ethics Framework and Standard of Conduct of the BC Public Service, was well as the Tri-Council Policy Statement 2, and the OECD Basic toolkit, among other sources.

Source: Behavioural Economics Team of the Australian Government. (2018[42]); Office of Evaluation Sciences. (2021[18]).

BOX 2.3.3: RISK ASSESSMENT IN APPLYING BEHAVIOURAL INSIGHTS FOR PUBLIC POLICY

The [Office of Evaluation Sciences](#) (OES) at the U.S. General Services Administration developed the Behavioural Insights Guide for Improving Payment Integrity which includes templates designed to illustrate the considerations involved in applying a behavioural approach to reduce improper payments. One of the included templates helps users prioritise the goals and objectives of their intervention while prompting them to consider the associated risks of their design.

Source: Office of Evaluation Sciences. (2021[18]). 'A Behavioral Insights Guide for Improving Payment Integrity. Retrieved April 7, 2022, from https://oes.gsa.gov/collaborations/improper-payments-playbook/

PRINCIPLE 2.4: PRESERVE FAIRNESS, EQUALITY, AND DIGNITY

Ethical BI for public policy aims to change a target behaviour, believed to better serve the interests of the public, derived from a legitimate policy mandate, and informed by deliberation with relevant stakeholders (cf. principle 2.2). However, this target behaviour can be addressed by a variety of potential interventions; some of which could create diverging effects among differing or similar populations. When targeting diverse populations, BI practitioners should strive to consider all possible options and opt for the intervention that is most effective at avoiding positive or negative discrimination (Lades & Delaney, 2020[13]). To do this, BI practitioners should:

Prompts 2.4:
- **Does the intervention promote discriminatory or offensive behaviours?**
- **Does your intervention stigmatise disadvantaged or marginalised populations?**
- **Did you take precautions to prevent unfairness in your sample population and in randomisation?**

PRACTICES 2.4:

- **Avoid promoting the uptake of discriminatory or harmful behaviours.** Positively changing the behaviour of one population may have a negative effect on another. Being aware of the potential spillover effects of your intervention can help preserve fairness and equality for all affected parties. This is also relevant to interventions that require action or resources from recipients in order for the outcome(s) to be realised – which in some cases may further exacerbate existing inequalities (see, Adams et al., 2016[43]; OECD, 2021[31]). Considering the degree of personal agency required of the target group to benefit from the intervention can help reduce creating new or reinforcing existing inequalities.

- **Avoid offending or stigmatising a subgroup of the population.** For example, cultural appropriations that inaccurately portray the religious or cultural practices of a certain population can be stigmatising and may do harm in falsely and negatively shaping others' opinions about that population (Lades & Delaney, 2020[13]).

- **Researchers may need to consider decompensation** designs and compensatory treatments to avoid unfair discrimination, as running a BI intervention could disadvantage a subpopulation as well as the control group (OECD, 2019[7]). Literature on sludge highlights the additional burdens individuals sometimes face when pursuing desirable choices, but also raises additional considerations on how BI sludges affect populations differently (Sunstein 2018[27], 2020[28]; 2021[44]; Thaler 2018[29]). This is why it is important to carefully consider whom your target audience is and to what extent they can access the intervention. For instance, an intervention requiring stable internet connection may cause those who do not have access to stable internet to be excluded from benefitting from the intervention entirely, and could result in further disadvantaging that population.

- **Take precautions to avoid using biased or flawed sources, datasets, and resources.** Just like the people who create them, databases, information systems, and platforms themselves are biased. Efforts to rigorously inspect the quality and validity of your sources and research can help avoid contaminating the results of your research by ensuring it is based on quality and trustworthy data.

- **Take precautions to prevent unfairness in randomisation.** Check that the experimentation does not induce significant inequalities and unfairness for participants or administers involved in the intervention. Ensure that at the end of the experiment, all groups are provided with the same information about the intervention, insofar as it does not affect the results of the experiment.

BOX 2.4.1: REDUCING LIMITATIONS OF RANDOMISATION

The Poverty Action Lab (J-PAL) has a toolkit intended for policymakers and practitioners applying randomised controlled trials (RCTs). The tool outlines common limitations and functional or ethical challenges in randomisation, real-world case studies, and solutions to such challenges (and their limitations) to support the use of experimentation in policymaking.

Source: Heard, K., O'Toole, E., Naimpally, R., & Bressler, L. (2017[46]). Real-World Challenges to Randomization and Their Solutions. United States: J-Paul North America. Retrieved from https://www.povertyactionlab.org/sites/default/files/research-resources/2017.04.14-Real-World-Challenges-to-Randomization-and-Their-Solutions.pdf

- **Anticipate possible inequitable impacts of policy interventions.** Some policy interventions might generate different distributional consequences, which need to be taken into consideration and carefully evaluated. For example, OECD work on food labelling in Chile highlighted how changes in the food environment may have different economic and health costs for different strata of society, as a similar basket of goods will carry a different weight on the income of different quintiles of the population (OECD, 2021[19]). As such, externalities and side effects of one measure on other aspects of citizens' life should always be evaluated, with a view of protecting the most vulnerable groups.

BOX 2.4.2: "BAN THE BOX" POLICY IN THE UNITED STATES

An increasing number of employers in the United States are adopting hiring policies that postpone the point in which applicants disclose their criminal records. The intention behind this policy is to reduce racial disparities and increase employment opportunities for applicants with criminal records. Despite its good intentions, Doleac and Hansen's (2020[45])research suggest that removing criminal history information from earlier stages of the application process appears to raise statistical discrimination of certain demographic groups. Their findings conclude that such policies reduced young, low-skilled, black men's probability of employment by more than 3 percentage points (5.1%) (Doleac & Hansen, 2020[45]). This is only one example of how interventions with good intentions may have unintended consequences that either neutralise the treatment or result in undesirable outcomes (more on this is in

3 RESEARCH AND EVALUATION
Gathering behavioural evidence responsibly

Applying behavioural science to public policy requires attention to scientific evaluation standards and evidence-based decision making whether it be desk research, systematic reviews, or conducting your own behavioural experimentation. Scientific standards and evidence-based design are common considerations during an intervention's design phase. This section promotes additional considerations for both before the design and after the implementation phases. Engaging in the below practices can help uphold good practices that safeguard the quality and integrity of BI in public policy while also promoting knowledge-sharing practices that serve the wider behavioural community working in government.

PRINCIPLE 3.1: ANTICIPATE AND PLAN FOR UNINTENDED CONSEQUENCES (SUCH AS BACKFIRE EFFECTS OR WELFARES LOSSES)

BI interventions are situated in the paradigm of evidence-based policy-making and generally involve pilot studies and rigorous testing before eventually being scaled. Compared to other policy approaches, behavioural testing reduces the risk of unexpected results once scaled or implemented. Nevertheless, BI interventions might not work out as planned and can result in unintended consequences. These may be first order backfire effects, caused when the intervention affects the target behaviour negatively, or second order backfire effects, occurring when the intervention causes divergent effects among different parts of the population. BI practitioners should take measures to prevent, detect, and address such effects regardless of if or when the intervention will be scaled. To do this, we recommend the following:

Prompts 3.1:
- **Have you referred to existing literature and experimentation that may help you anticipate and address, in advance, unintended results or consequences?**
- **Have you established procedures or protocols to initiate if negative consequences arise?**

PRACTICES 3.1:

- **Reflect on different scenarios with a diverse and representative sample of stakeholders** to mitigate the possibility of unintended consequences from the intervention. Facilitate opportunities for discussion about the contextual conditions that may alter the effectiveness of the results, such as social, cultural, demographic, political, legal, and economic considerations.

- **Consider running small-scale or pilot studies where appropriate.** Pre- and small-scale testing can help provide preliminary insights that are useful for determining whether to proceed with larger testing samples and identify potential challenges or consequences with a given intervention design (for more on conducting pilot studies, see Government Chief Social Researcher's Office, 2003[47]; Lunn & Robertson, 2018[48]; Pearson et al., 2020[49]).

- **Plan for experiments to be paused or terminated** if negative side effects unexpectedly occur or if one experimental group experiences dramatically worse results than another. Aim to establish these determinates with partner institutions as early in the design phases as possible to avoid instances where such decisions may be subjected to a higher degree of subjectivity and dispute. Planning fixed and reoccurring check-ins can help ensure ethics are being considered, monitored, and assessed throughout the project, and ensure appropriate suspension or termination of the experiment if ethical concerns should arise.

BOX 3.1.1: IN CASE FRAMEWORK FROM UK'S GOVERNMENT COMMUNICATION SERVICE

The United Kingdom's Cabinet Office Government Communication Service Behavioural Team has developed the IN CASE framework to help anticipate intended and unintended behaviour changes triggered by their communication campaigns by prompting the following considerations:

- **I**ntended behaviour

- **N**on-target audiences

- **C**ompensatory behaviours

- **A**dditional behaviours

- **S**ignalling

- **E**motional impact

The framework invites an exploration of the consequences of encouraging the intended behaviour, the possible response and effects on non-target groups, the compensatory and additional behaviours adopted as consequences of the behaviour change, signalling effects and the emotional impact of the campaign. This allows for identifying unintended behavioural consequences of communication campaigns and offers ways to mitigate these risks.

Source: UK Government Communication Service. (2021[50]). IN CASE: A behavioural approach to anticipating unintended consequences. United Kingdom: Government of United Kingdom. Retrieved April 7, 2022, from https://gcs.civilservice.gov.uk/publications/in-case-a-behavioural-approach-to-anticipating-unintended-consequences/.

PRINCIPLE 3.2 PRE-REGISTER: DEFINE RESEARCH QUESTIONS, HYPOTHESES, AND METHODS BEFORE OBSERVING THE OUTCOMES

Pre-registration is an increasingly common practice among researchers. However, its benefits also extend to the BI community of practitioners. The advantages of pre-registration include addressing publication bias by acknowledging the value of both intended and unintended outcomes, including null results. It also contributes to practices of transparency and openness by avoiding ad hoc misinterpretations of the results and contributing to the growing body of BI research. This opens behavioural science research to the possibility of pre-analysis feedback and review from other researchers and the BI community.

Prompts 3.2:
- **Have you pre-registered your research, including observational and experimental research?**

BOX 3.2.1: OECD'S PRE-REGISTRATION PORTAL

In 2021, the OECD launched a pre-registration portal dedicated and adapted to BI practitioners, allowing them to publish the essential information about their different research projects and trials, even before starting projects. Importantly, the OECD's pre-registration portal is a dedicated registry to pre-register and publish all projects (not only experimental projects). It serves as a mechanism to set standards for openness and rigor by making projects discoverable early in their lifecycle and making pre-registration easy for BI practitioners.

One may choose to include general characteristics of the project, the methods used (experimental and others), start date, hypothesis, design, analyses, and outcome variables are required. The portal includes an embargo option, meaning users can include information but keep it hidden from public access until a predetermined date. Practitioners also have the option to include additional information, as well as to upload their pre-analysis plan and document research outcomes after pre-registration.

The portal is a particularly useful resource for the public by offering a place to discover ongoing projects and navigate case studies in real time within the connected OECD infrastructure.

Source: OECD. (2021[51]). BI Pre-registration form. Observatory of Public Sector Innovation.
https://oecd-opsi.org/bi-pre-registration-form/

Moreover, pre-registration promotes transparency in the policy-making process and accountability to the public. As mentioned in Section 2.1, we acknowledge that BI practitioners rarely have a say in what information is approved for publication. However, BI practitioners are familiar with the many benefits of publishing data, particularly for scientific rigor and knowledge-sharing, making them uniquely positioned to advocate for pre-registration. As such, when possible, BI practitioners should:

PRACTICES 3.2:

- **Define the research questions, hypotheses, methods, and analysis plan before observing the research outcomes** and, if possible, register them on a dedicated platform whether academic, governmental, or other (see Box 3.2.1 for example).

- **If you decide to run additional analyses include this in the report** and share which exploratory analyses you plan to run.

BOX 3.2.2: PRE-REGISTERING TRIALS AT OES AND BETA

The United States' OES as well as Australia's BETA have begun pre-registering all trials and analysis plans. This allows readers to track any changes made from the originally planned analysis to the ones eventually ran and reported, that are duly justified as post-commitment adjustments. Thus, the interpretation of data is strictly constrained to a priori and post hoc analysis (that are run only after observing the data and not based on the hypothesis and research question, and thus are suspicious of manipulation to draw a biased conclusion) are clearly identified as such and as only exploratory.

OES pre-analysis plans include a project description, hypothesis, data description, independent and outcome variables, data exclusion and missing data treatment, and all planned data analysis (descriptive, hypothesis testing and follow-up).

BETA publishes some trial information on its own website in addition to pre-registering its analysis plan, such as the trial start and end date, ethics approval, experimental design, outcome(s), expected sample size, as well as links to their pre-registration reports.

Source: Behavioural Economics Team of the Australian Government. (2022[52]); Office of Evaluation Sciences. (n.d.[53]).

PRINCIPLE 3.3: PROTECT DATA, PRIVACY, CONFIDENTIALITY, AND OBTAIN INFORMED CONSENT WHEN NECESSARY

Gathering data to generate insights about human behaviour raises specific ethical challenges, regardless of the collection method or reason(s) for use. Analysing behaviour, as well as testing and evaluating policy interventions requires collecting and using data, including personal data, to generate insights or to generalise populations. Furthermore, it these insights are used to inform policy outcome(s) that may or may not encourage changes in individuals' behaviours, prompting additional ethical considerations. Deliberations for ensuring the collection of ethical, timely and informed consent should include the below practices:

Prompts 3.3:
- **Have you gathered the necessary consent from participants that preserves their autonomy?**
- **Did you disclose all information possible to inform participants while preserving the integrity and outcome(s) of the intervention?**
- **Are there sufficient measures in place that prevent the misuse of confidential information and data?**

PRACTICES 3.3:

- **When possible, clearly inform participants** about the research and its aims. Participants should be given as much information as possible pertaining to the process, purpose, outcome(s), and implication(s) of the intervention at the appropriate times to promote informed consent. A point-of-contact should be established to monitor feedback channels and respond to follow-up questions about processes, procedures, and information relating to the intervention.

- **Ensure participation is voluntary.** Participants should reserve the right to opt-out of or refuse participation in the intervention at any time without fear of punishment or negative repercussions. In cases where involvement is involuntary, interventions should be easily avoidable (cf. principle 2.1).

- **Collect appropriate consent** from guardians, parents, or legal representatives in situations where participants are unable to provide consent for themselves (such as minors), whenever necessary. Ensure this consent is properly collected, documented and stored such that it cannot be accessed by those without the proper authority. For more on the specific aspects of consent collection and storage, see The American Psychological Association's Ethical Principles of Psychologists and Code of Conduct (2003[41]).

- When it is not possible to collect consent, **pay extra caution towards vulnerable groups and consider additional measures to enhance accountability** such as seeking independent ethic board approval in lieu of consent from participants or notifying and gathering consent from participants in a follow up phases of testing (cf. box 2.3.1).

- **Protect personal data, privacy, and confidentiality.** Behavioural science draws on data to generate useful insights about the preferences and behaviours of individuals. Using this data responsibly means protecting users' personal information and data by establishing protocols to safeguard personal data from abuse or misuse. Consider using procedures that ensure the confidentiality of participants (e.g., randomised response methods) or choose not to collect or connect any data to identifiers. Physical and digital safety measures such as locked file folders, strong passwords, and encrypted digital folders should be used to prevent data breaches and unauthorised data access. When considering the ethical concerns specific to collecting and using personal data for the public sector, you can refer to the OECD's Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data (OECD, 2013[54]) and the practices set out in the Good Practice Principles for Data Ethics in the Public Sector (OECD, 2021[12]).

BOX 3.3.1: RANDOMIZED CONTROLLED TRIALS WITHOUT CONSENT

In a systematic review of ten clinical research proposals applying randomisation without consent (RWOC), Flory et al. (2016[55]) find that most designs that do so raise ethical questions by not appropriately respecting patient's autonomy. They find only two scenarios whereby RWOC may be justified – cluster randomized designs and emergency research (which typically applies more appropriately to medical research). For behavioural insights, the generalisability of the results, or instances where the Hawthorn Effect1 may skew the data, resulting in flawed data. In such cases, collecting informed consent may not be possible.

Much like applying behavioural approaches in the first place, any experimentation that chooses to proceed without participant consent should establish a strong case for doing so. Consult with relevant literature and policies to gain further insights into whether your justification for experimentation without consent is valid and what protocols to follow in these cases.

BOX 3.3.2: USING BEHAVIOURAL INSIGHTS TO IMPROVE CONSENT GATHERING IN GREATER MANCHESTER

If reducing burden on the participants and researchers is the goal, consider applying BI to reduce the cognitive load on those participating in BI experimentation by simplifying the consent process for both researchers and participants in a way that respects participants' autonomy.

The greater Manchester Combined Authority employed behavioural science to alter the way their Privacy Notice was presented to participants. One rewrite simply reduced the amount of information presented to participants. The second rewrite simplified the notice to key messages and featured visual components like graphs and images. Both rewrites were found to increase comprehension compared to the pre-existing notice and neither appeared to affect participants' willingness to consent to the Privacy Notice.

Source: Local Government Association of Greater Manchester. (2017[56]). 'Using behavioural insights to improve complex consent processes'. Retrieved March 14, 2022, from https://www.local.gov.uk/using-behavioural-insights-improve-complex-consent-processes-greater-manchester-combined-authority

PRINCIPLE 3.4: PUBLISH AND SHARE RESULTS

Similar to arguments for pre-registering your project, publishing all results, including negative and null results, advances research and policy-making processes and contributes to the exchange of knowledge and best practices among BI practitioners and experts. The following should be considered regarding sharing your intervention results:

Prompts 3.4:
- **Are you and your partner institution(s) prepared and willing to share all results, including negative and null results?**
- **Are you inaccurately interpreting or mis-representing data?**
- **When possible, do you publish your research, including observational and experimental research?**

PRACTICES 3.4:

- Efforts should be made to **ensure data and results are portrayed as accurately as possible.** Instances in which data or results are misrepresented for any reason should be avoided. In cases in which representation of findings were incorrect, notify the relevant stakeholders to prevent the error from negatively affecting the outcome(s) of the project or others' interpretation or understanding of the data.

- **To whatever extent possible, results should be shared publicly,** even if they are null, do not confirm the hypotheses, or cannot generate important policy recommendations.

- **Non-significant and negative results help the BI community** to learn just as much as significant and positive ones. Reporting thes as well as challenges and barriers publicly (such as government websites or academic journals) when possible, increases the spread of knowledge and best practices and may bring ethical clarity for others running similar tests or working in relevant areas of interest (Evans, 2018[58]; Office of Evaluation Sciences, n.d.[59]).

- If possible and **under the condition of anonymisation, data should be made publicly available**, and offered as open data to allow replication of results and encourage accessibility to data. When data is anonymised such that personal identifiers (i.e., name, address, date of birth, etc.) cannot be traced to the participant, open data can be useful for identifying economic, social, and cultural trends that can inform policy decisions (OECD, 2021[12]).

BOX 3.4.1: GOOD PUBLISHING PRACTICES

Since 2016, BETA has committed to publishing the results of their trials. This practice creates an expectation within the team, and with external partners, that results of trials will be disseminated, regardless of the results, signalling their commitment to openness. BETA's website lists the details of pre-registered trials and completed trials. Once trials are completed, a full report is published on the website alongside a one-page summary. The team publishes technical appendices and statistical tables to allow for scientific scrutiny.

Similarly, the United States' OES has a commitment to publish various aspects of a project's lifecycle from project initiation to impact evaluation. This guide designed by the OES outlines the steps to make results publicly available, which can be used across a wide variety of testing methodologies and policy areas.

Source: Behavioural Economics Team of the Australian Government. (2022[52]).; Office of Evaluation Sciences. (n.d.[60]). OES Project Process. Retrieved April 11, 2022, from https://oes.gsa.gov/projectprocess/

BOX 3.4.2: OECD'S BI KNOWLEDGE REPOSITORY

Along with the pre-registration portal (cf. box 3.2.1), the OECD launched a Global BI Project Repository, allowing for the BI community to submit completed projects into a database accessible and updated by BI teams from around the world. A minimal amount of information is required to make sharing easy: general key features, methods, design, and results; additional information can be added, such as final reports and findings. The tool allows projects to be classified by unit, policy area, behavioural tool, among others, making the database easy to navigate.

Source: OECD. (2021[61]). BI Projects Archive. https://oecd-opsi.org/bi-projects/

4 POLICY IMPLEMENTATION

Preserving ethics when implementing and scaling BI results
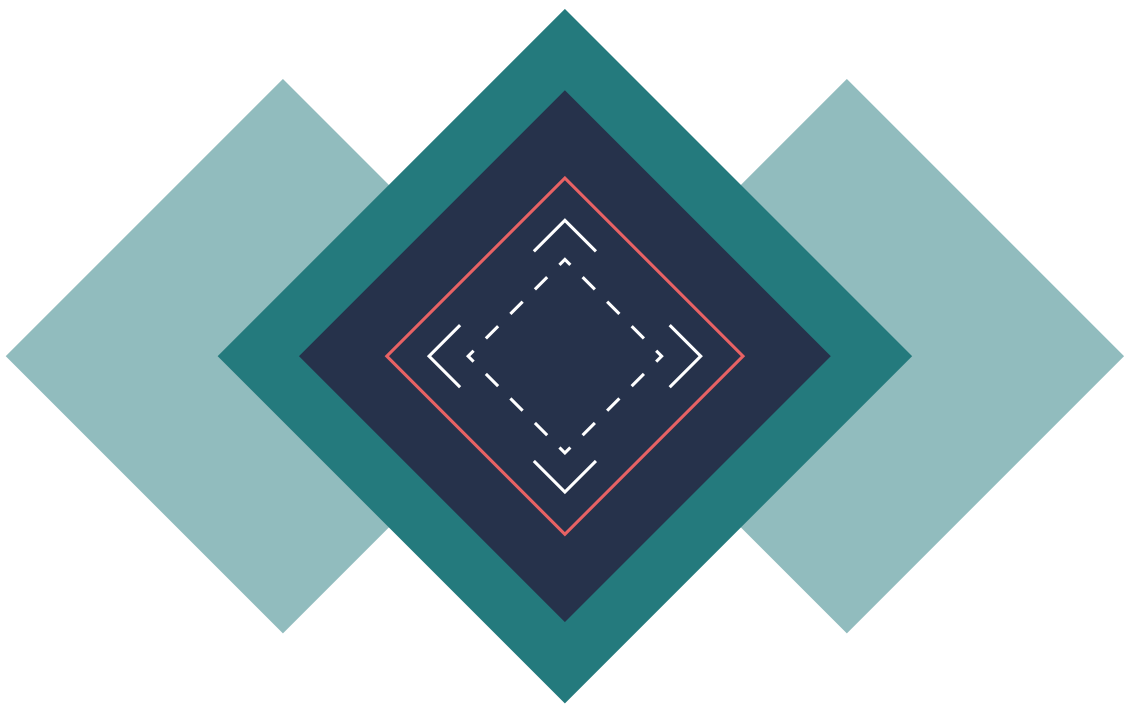
Ethically designing and testing an intervention does not guarantee that the intervention will remain ethical as a scaled policy. For instance, effects demonstrated on a sample might not reproduce on an entire targeted population and in some cases, may even cause unintended and harmful side-effects that render the policy ineffective or unethical. The recommendations put forward in this section are intended to minimise these risks by acknowledging the limitations of experimentation when scaling an intervention.

PRINCIPLE 4.1: ENSURE ETHICAL CONTINUITY WHEN ADAPTING AND SCALING

BI practitioners might arrive at the stage of implementation by two distinct pathways: either via a problem-oriented approach, meaning one has diagnosed a problem, assessed intervention options, and tested these options themselves, or via a solution-oriented approach, meaning they, or policy makers, want to adapt an already successfully tested BI intervention.

Choosing a solution-oriented approach offers useful opportunities to learn from others' experience but requires additional analysis of their relevance and applicability in the context to which you propose their application. Beyond assessing whether results are statistically significant, or possess external validity, BI practitioners must also evaluate if the results are reliable, credible, and meaningful. Assume for example, an RCT in the US shows that making vaccine appointments for everyone by default boosts vaccination rates. Would it be equally effective and ethical for the South African government to introduce default appointments also?

Prompt 4.1:
- **Have you considered new ethical concerns resulting from scaling and adapting in new contexts?**

There are many conditions, such as political, legal, cultural, technological, and economic factors, that influence both the probability of a policy being implemented and its impact once implemented. When adapting the results of an intervention from one environment into another, it is important to understand the factors that contributed to the success of the intervention and identifying whether the same conditions exist or need to exist in a new or altered environment in order for the intervention to be successful. To assess this and similar questions, consider the following questions adapted from Hallsworth and Kirkman:

- » Is the sample large and representative enough?
- » Are the groups comparable?
- » Is the size of the effect similar or comparable to what we see in similar research?
- » Do the outcome(s) include all data? If not, why not?
- » Do the outcome(s) truly measure what they claim to?
- » Is the study sufficiently powered to detect realistic effect sizes in all primary outcomes?

» Is the study sufficiently designed that a null result is informative?

» Does an outcome of this size mean those who experience it benefit in a meaningful or significant way?

» Do the short, medium, and long-term effects of the intervention support a continued investment in this intervention over others?

» What are the other conditions, beyond intervention design, that are required for the success of this intervention (i.e., cultural, social, legal, political, economic, social, and technological)?

» Does the adapted environment provide the necessary conditions to indicate that the intervention can be successfully scaled?

These questions are most relevant for a solution-oriented approach but may also be useful to ask regarding one's own research to adopt a critical perspective.

In both approaches, BI practitioners need to assess whether the results of their study possess external validity and can be generalisable and applied to a larger sample or an entire population. Considering one's policy context also means considering additional conditions beyond who is or is not targeted. For example, one may have a representative sample, yet running an experiment before the 2019 coronavirus disease (COVID-19) pandemic may generate results that cannot be replicated in a post-pandemic scenario. To assess related questions, consider the following:
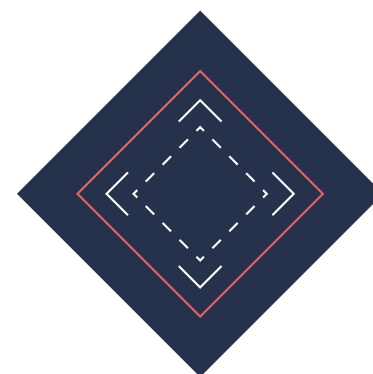
PRACTICES 4.1:

- **Review results from relevant experiments and research and (re)assess their quality.** Assess if some side effects, detrimental effects on sub-samples, backfire effects, or any other unintended consequences emerged through implementation and/or scaling. Plan for how to manage these effects and how to avoid them in your implementation. Indeed, discriminatory effects observed from testing may not ethically justify scaling, even if sub-target results could be achieved (cf. boxes 4.1.1, 4.1.2, and 4.1.3).

- **Reassess ethical questions against the adapted contexts and populations.** Gather data about potential implications on new target population and unintended populations and the environment of the intervention. Consider redrafting original (pre)analysis plans and processes and consider reviewing Sections 1 and 2 of this guide to ensure ethical continuity when intervention is adapted into a new context.

- **Identify and record changes** in the intervention between the project for which evidence has been gathered and the one that will be implemented or scaled (i.e., changes in methodology, target group, analysis and models, relevant stakeholders, etc.) Be sure to remain as close as possible to the design that was empirically tested. Anticipate conditions of the new context that will require adaptation of the intervention and provide guidelines for the efficacy and ethical principles to be maintained or improved to the relevant actors.

- **Respect the relevant political institutions and their mandates** when incorporating BI research and experimentation into the policy-making cycle, to ensure the proper executive and legislative oversight of the project and its outcomes.

BOX 4.1.1: STUDY STUDENT FINANCIAL AID IN THE UNITED STATES

A 2020 report discussed a campaign that tested different encouragement messages to get 800,000 pre-college high schoolers to take advantage of federal student financial aid in the United States (Bird et al., 2021[62]). Several variations that had been proven to be effective in previous local studies were tested again in different environments. When scaled, none of the messages proved to be effective in boosting neither the receipt of financial aid nor college enrolment.

The authors' reflections do not suggest that the previous studies are flawed, but rather, they hypothesise that some features of local campaigns may have contributed more to the original effects than previously understood. For example, students may feel more connection with a local organisation sending messages or may perceive such messages to be more personal than when they are part of a national campaign. While the exact reasons are unknown, the idea remains clear: Do not assume that results will be sustained without first understanding the conditions that led to them in the first place.

Source: Bird, K. A., Castleman, B. L., Denning, J. T., Goodman, J., Lamberton, C., & Rosinger, K. O. (2021[62]). Nudging at scale: Experimental evidence from FAFSA completion campaigns. Journal of Economic Behavior and Organization, 183, 105–128. https://

BOX 4.1.2: BACKFIRE EFFECTS IN ENERGY USE

One example of a backfire effect occurred in an intervention that leveraged social norms to reduce energy consumption in homes. A 2020 study set out to understand the impact of social norms on household energy consumption. The results found social norms to be effective on the households consuming the most energy noting that they were well above the average. Conversely, the impact of social norms was found less effective on the households that were consuming less-than average amounts of energy because they were confronted with how little energy they use compared to other homes and therefore, would need to increase their energy use in order to comply with the average household consumption of energy.

This example demonstrates the importance of understanding the target audience and the environment in which you employ behavioural levers. Failing to accurately account for these may render your intervention ineffective or may cause the opposite of the desired outcome.

Source: Bonan, J., Cattaneo, C., d'Adda, G., & Tavoni, M. (2020[63]). The interaction of descriptive and injunctive social norms in promoting energy conservation. Nature Energy, 5(11), 900–909. https://doi.org/10.1038/s41560-020-00719-z

BOX 4.1.3: SAVING RATES IN MEXICO

An intervention in Mexico tried to increase savings rates by sending different SMS reminders to citizens (Shah et al., 2018[64]). The analysis revealed that the reminder with the largest effect size actually made women under the age of 29 much less likely to contribute to their savings compared to other interventions. In contrast, individuals aged 29 to 41 years increased their savings after receiving the same message.

Many countries have laws that guarantee individuals equal treatment from the public and private sector. If we find, as in the Mexico savings study, that women under 29 respond differently from the general population, then behavioural scientists and policy makers should aim to understand what measures are required to preserve the fair treatment of individuals when subjected to behavioural interventions and avoid unintended consequences during scaling that negatively affect one group over another.

Source: Shah, A., Osborne, M., Lefkowitz, J., Fertig, A., Soman, D., & Mazar, N. (2018[64]). Can Making Family Salient Improve Retirement Contributions? Evidence From Field Experiments in Mexico. ACR North American Advances, NA-46. https://www.acrwebsite.org/volumes/2411999/volumes/v46/NA-46.

PRINCIPLE 4.2: ENSURE COMMUNICATION AND IMPLEMENTATION GUIDANCE WITH PARTNER INSTITUTION(S)

BI practitioners not only have the responsibility to conduct safe and fair testing, but they also have a pedagogical responsibility towards policy makers and other practitioners. However, this can be challenging to fulfil when BI practitioners are left out of the implementation and scaling phases of an intervention, despite possessing insightful knowledge about the conditions that contribute to the success and failure of interventions. As such, efforts should be made by government officials to include BI practitioners in all phases of the intervention cycle. Similarly, efforts should be made by BI practitioners to help policy makers understand and interpret implications and limitations of the insights collected. Conscious that the effect of the intervention in a controlled environment may differ greatly from its real-world application, BI practitioners should aim to equip partner institutions, whether academic, public, or private, with the appropriate guidance or best practices for implementing and scaling behaviourally-informed interventions. Although that is rare the practitioners will have the final say on whether an intervention is scaled or not, their ability to link the outcomes of an experiment with policy goals are a vital component to the success of the behavioural project.

BOX 4.2.1: SCALE UP TOOLKIT

BehaviourWorks Australia and the Victorian Government Behavioural Insights Unit have developed an evidence-informed toolkit to help BI researchers and practitioners to start behavioural interventions with scaling in mind, including how to:

- Learn about scaling policies, its challenges, and useful frameworks;
- Assess the feasibility of different intervention ideas;
- Select a scalable behaviour change intervention;
- Design or adapt an intervention for testing and scale up; and
- Test scale up assumptions about your intervention in a pilot or trial.

Each tool features free instructional videos and templates available to anyone. The toolkit's broad design ensures that it can be easily applied by anyone looking to scale BI interventions, regardless of their educational or professional background.

Source: BehaviourWorks Australia. (2020[65]). Scale up toolkit. https://sites.google.com/monash.edu/behaviourworks-scaleup-toolkit/home; Saeri, A. K., Slattery, P., Tear, M. J., Varazzani, C., Epstein, D., Knott, C., Kusmanoff, A., Knott, C., Bagshaw, H., Phillips, K., & Liao, J. (2021[66]). Scale up of behaviour change interventions: A rapid review of evidence and practice. OSF Preprints. https://osf.io/download/60077d78ba010908a2891f4d/

Together, government officials and BI practitioners can help prevent work from being distorted (e.g., generalising across different contexts) or misinterpreted (e.g., taking non-significant findings as positive evidence) and enhancing the quality and efficacy of behaviourally-informed policy Lades & Delaney, 2020[13]). For this, consider the following:

Prompts 4.2:
- **Have you provided partner institution(s) with advice and/or guidelines for implementation and scaling?**

PRACTICES 4.2:

- When possible, **inform officials that will oversee the program on the interventions** and how it has been tested in order to uphold ethical consistency during implementation.

- During each stage of the intervention process, **be as open and detailed as possible** with your partner institution(s) to increase awareness and knowledge about the design, objectives and effects, and potential challenges and implications of the intervention.

- BI practitioners should **strive for post-experimentation follow-up and communication** that offers guidance in the implementation and scaling process. This can help with the transition from testing to application phases and reduce instances where the results of the intervention are misinterpreted or ignored for unjustified reasons.

PRINCIPLE 4.3: BE ACCOUNTABLE AND ACCESSIBLE TO THE PUBLIC

It is no surprise that BI provides key insights into the way critical information is disseminated and accepted by individuals (OECD, 2021[67]). Using behavioural levers to improve the ways in which the intentions, purposes, outcome(s) and result(s) of your intervention are communicated to the public can help ensure the right information is reaching the right people. In doing so, BI teams and their partner(s) can increase their openness and accountability to the public they serve, while reinforcing the impact of behavioural science in public policy in the process. Consider extending your applications of behavioural science and use BI to inform the way you communicate your findings with the public. Aim to design simple and visually appealing material that are approachable and easily digestible to the masses (Service et al., 2021[68]; Dutch Ministry of General Affairs, 2021[24]). Public engagement is necessary in preserving the integrity and credibility of behavioural science, both as a research methodology and as a policy-making tool. As such, strive for accountability and accessibility when and where possible through the practices below:

Prompts 4.3:
- **When possible, do you actively communicate with citizens through feedback channels?**
- **Are the feedback channels accessible and managed in a timely manner?**

PRACTICES 4.3:

- **Encourage active communication** on the policy and the overall process that informed it in an accessible and approachable manner. Consider the general public as your main audience and simplify messages, reducing technical jargon, and highlight key takeaways to make your research more approachable to non-practitioners (for more details of public communication methods and approaches (see the OECD Report on Public Communication, 2021[67]). Consider having a separate section containing scientific language to encourage rigor and peer-review of methodologies and results. Strive to make this easily accessible to the public, especially those most likely to be affected by the intervention.

- **Document and share a digestible report for all projects** (including exploratory ones) which lays out all ethical and methodological considerations (such as goal, intervention, research, ethics, board approval etc.) and explains the intervention and research design. Ensure to properly associate and mention the relevant institutions that run or contribute to the experiments (see box 4.3.2 for examples).

- **Seek opportunities to share insights and feedback with participants.** Even without disclosing confidential information about the study, relaying key insights and the overall results of an intervention with participants may increase participants' overall satisfaction with enrolling in the intervention. This allows for participants to familiarise themselves with research and policy-making processes while also reminding participants that their time and participation is valuable and contributes to useful results (see box 4.3.1 for example).

- **Perform follow-up analyses to evaluate performance and quality of research.** Follow-up analyses are an effective way to reflect on the challenges and successes of BI experimentation following its completion. This can be done with your team however, including others who were involved in the project can offer valuable insights for yourself and others, for establishing best practices, identifying opportunities for improvement, and gathering feedback for future consideration, including the ethical quality of your work.

- **Advocate for the existence of feedback channels,** both for people impacted by the project and the general public who may have comments and questions regarding the activity of the unit.

BOX 4.3.1: BUSARA'S FOLLOW-UP COMMUNICATIONS WITH PARTICIPANTS

After publishing the results of a 2021 study, the Busara Center for Behavioural Economics sent out SMS messages to all participants of their qualitative sample as a cost-effective way of establishing follow-up communication with participants. The messages were followed by phone calls whereby participants were able to provide researchers will feedback about the way results were communicated and disseminated, trustworthiness and comfort in participating in the experiment, and whether they found the overall experience and results valuable.

Participants expressed interest in learning more about the impact and implications of the study as a means of sharing their experiences and knowledge with others like friends and family and found feedback channels as an effective way to do this.

Source: Wein, T., Schilling, M., Hammond, P., Mumo, J., & Juma, C. (2021[69]). Value and validation. How feedback enhances the quality of research outputs. The Busara Blog. https://medium.com/busara-center-blog/value-and-validation-113750e7c0ad

BOX 4.3.2: ACCESSIBLE AND APPROACHABLE REPORT SUMMARIES DESIGNED FOR CITIZENS

BETA publishes one-page summaries designed for citizens rather than policy makers. This aims to make their work more transparent and accessible to the public. Similarly, the Joint Research Committee (JRC) of the European Commission is increasing accessibility of their publications by including brief summaries of their findings in clear and simplified formats. Their summaries are available through the repository, where users can access the file directly by downloading it or scanning a QR code.

Source: Behavioural Economics Team of the Australian Government,(2022[52]); Joint Research Centre,(2022[70]).

REFERENCES

43   Adams, J., Mytton, O., White, M., & Monsivais, P. (2016). Why Are Some Population Interventions for Diet and Obesity More Equitable and Effective Than Others? The Role of Individual Agency. PLOS Medicine, 13(4), e1001990. https://doi.org/10.1371/journal.pmed.1001990

3   Barton, A., & Grüne-Yanoff, T. (2015). From Libertarian Paternalism to Nudging—And Beyond. Review of Philosophy and Psychology, 6(3), 341–359. https://doi.org/10.1007/s13164-015-0268-x

17   Behavioural Economics Team of the Australian Government. (n.d.). Behaviour Discovery Tool. Retrieved April 7, 2022, from https://youropinion.au1.qualtrics.com/jfe/form/SV_bJBe80XcOFtpHjD

42   Behavioural Economics Team of the Australian Government. (2018). BETA Ethical Risk Assessment Form. https://www.pmc.gov.au/sites/default/files/foi-log/foi-2019-287-docs-for-release.pdf

52   Behavioural Economics Team of the Australian Government. (2022). Projects. Retrieved April 11, 2022, from https://behaviouraleconomics.pmc.gov.au/projects

21   Behavioural Insights Team. (n.d.). Barrier Identification Tool. BIT Barrier Tool. Retrieved March 25, 2022, from https://www.bitbarriertool.com

65   BehaviourWorks Australia. (2020). Scale up toolkit. https://sites.google.com/monash.edu/behaviourworks-scaleup-toolkit/home

14   Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W. J., & Galing, S. (2017). Should Governments Invest More in Nudging? Psychological Science, 28(8), 1041–1055. https://doi.org/10.1177/0956797617702501

62   Bird, K. A., Castleman, B. L., Denning, J. T., Goodman, J., Lamberton, C., & Rosinger, K. O. (2021). Nudging at scale: Experimental evidence from FAFSA completion campaigns. Journal of Economic Behavior and Organization, 183, 105–128. https://doi.org/10.1016/j.jebo.2020.12.022

63   Bonan, J., Cattaneo, C., d'Adda, G., & Tavoni, M. (2020). The interaction of descriptive and injunctive social norms in promoting energy conservation. Nature Energy, 5(11), 900–909. https://doi.org/10.1038/s41560-020-00719-z

1   Bonde, S., & Firenze, P. (2013). A Framework for Making Ethical Decisions. Brown University. https://www.brown.edu/academics/science-and-technology-studies/framework-making-ethical-decisions

36   Bramhall, S. (2011). Presumed consent for organ donation: A case against. Annals of The Royal College of Surgeons of England, 93(4), 270–272. https://doi.org/10.1308/147870811X571136b

35   Bruns, H. (2021). Nudges can be transparent and yet effective. European Commission. https://publications.jrc.ec.europa.eu/repository/handle/JRC125313#:~:text=Nudges%2C%20policy%20instruments%20that%20seek,reservations%2C%20but%20is%20rarely%20considered.

2   Costa, E., & Halpern, D. (2019). The behavioural science of online harm and manipulation, and what to do about it. The Behavioural Insights Team. https://www.bi.team/wp-content/uploads/2019/04/BIT_The-behavioural-science-of-online-harm-and-manipulation-and-what-to-do-about-it_Single.pdf

39   Devine, H. (2021, August). OECD Interview with Behavioural Insights Experts Working in Government [In-person interview].

20   Dolan, P., Hallsworth, M., Halpern, D., King, D., & Vlaev, I. (2010). MINDSPACE. UK Cabinet Office. https://www.bi.team/publications/mindspace/

45   Doleac, J. L., & Hansen, B. (2020). The Unintended Consequences of "Ban the Box": Statistical Discrimination and Employment Outcomes When Criminal Histories Are Hidden. Society of Labor Economists, Economics Research Center, 38(2). https://doi.org/10.1086/705880

24   Dutch Ministry of General Affairs. (2021). CASI – Communication Activation Strategy Instrument. Dutch Ministry of General Affairs. https://www.communicatierijk.nl/documenten/publicaties/2021/12/09/casi-in-english#:~:text=The%20Communication%20Activation%20Strategy%20Instrument,building%20blocks%20for%20your%20interventions.

58   Evans, D. (2018). How to Publish Statistically Insignificant Results in Economics. World Bank Blogs. https://blogs.worldbank.org/impactevaluations/how-publish-statistically-insignificant-results-economics

55   Flory, J. H., Mushlin, A. I., & Goodman, Z. I. (2016). Proposals to Conduct Randomized Controlled Trials Without Informed Consent: A Narrative Review. Journal of General Internal Medicine, 31(12), 1511–1518. https://doi.org/10.1007/s11606-016-3780-5

47   Government Chief Social Researcher's Office. (2003). Trying It Out: The role of "pilots" in policy-making. UK Cabinet Office. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/498256/Trying_it_out_the_role_of_pilots_in_policy.pdf

16   Hallsworth, M., & Kirkman, E. (2020). Behavioral Insights. MIT Press. https://mitpress.mit.edu/books/behavioral-insights

15   Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. Journal of Public Economics, 148, 14–31. https://doi.org/10.1016/j.jpubeco.2017.02.003

5   Hansen, P. G., & Jespersen, A. M. (2013). Nudge and the Manipulation of Choice: A Framework for the Responsible Use of the Nudge Approach to Behaviour Change in Public Policy. European Journal of Risk Regulation, 4(1), 3–28. https://doi.org/10.1017/S1867299X00002762

46   Heard, K., O'Toole, E., Naimpally, R., & Bressler, L. (2017). Real-World Challenges to Randomization and Their Solutions. J-Paul North America. https://www.povertyactionlab.org/sites/default/files/research-resources/2017.04.14-Real-World-Challenges-to-Randomization-and-Their-Solutions.pdf

70   Joint Research Centre. (2022). JRC Publications Repository. JRC Publications Repository. https://publications.jrc.ec.europa.eu/repository/

34   Kahneman, D. (2011). Thinking, fast and slow. Straus and Giroux.

13   Lades, L., & Delaney, L. (2020). Nudge FORGOOD. Behavioural Public Policy. https://www.cambridge.org/core/journals/behavioural-public-policy/article/nudge-forgood/06BC9E9032521954E8325798390A998A

56   Local Government Association of Greater Manchester. (2017). Using behavioural insights to improve complex consent processes. https://www.local.gov.uk/using-behavioural-insights-improve-complex-consent-processes-greater-manchester-combined-authority

48   Lunn, P., & Robertson, D. (2018). Using Behavioural Experiments to Pre-Test Policy. Economic & Social Research Institute.

57   Mahtani, K., & Spencer, E. (2017). Sackett Catalogue of Bias. Catalogue Of Bias. https://catalogofbias.org/biases/hawthorne-effect/

25   Michie, S., van Stralen, M. M., & West, R. (2011). The Behaviour Change Wheel: A new method for characterising and designing behaviour change interventions. PubMed, 6(1), 42. https://doi.org/10.1186/1748-5908-6-42

54   OECD. (2013). Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data. OECD Legal Instruments. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0188

8   OECD. (2017). Recommendation of the Council on Open Government. OECD Legal Instruments; OECD/LEGAL/0438. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0438

32   OECD. (2017). Use of Behavioural Insights in Consumer Policy (No. 36). OECD Publishing. https://www.oecd-ilibrary.org/docserver/c2203c35-en.pdf?expires=1650453705&id=id&accname=guest&checksum=48C663F152099E90869EF4A836A0FDAA

26   OECD. (2019). Assessing incentives to reduce traffic congestion in Israel. OECD Publishing. https://www.oecd.org/economy/israel-economic-snapshot/

33   OECD. (2019). Delivering Better Policies Through Behavioural Insights. OECD Publishing. https://doi.org/10.1787/6c9291e2-en

10   OECD. (2019). Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies | en | OECD. OECD Publishing. https://www.oecd.org/publications/enhancing-access-to-and-sharing-of-data-276aaca8-en.htm

9   OECD. (2019). OECD Principles for Artificial Intelligence. OECD Legal Instruments. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

7   OECD. (2019). Tools and Ethics for Applied Behavioural Insights: The BASIC Toolkit. OECD Publishing. https://www.oecd.org/gov/regulatory-policy/tools-and-ethics-for-applied-behavioural-insights-the-basic-toolkit-9ea76a8f-en.htm

30   OECD. (2020). Consumer policy in the digital transformation. In OECD Digital Economy Outlook 2020. OECD Publications. https://www.oecd-ilibrary.org/science-and-technology/oecd-digital-economy-outlook-2020_bb167041-en

23   OECD. (2021). Behavioural Insights for Better Tax Administration: A Brief Guide (Forum on Tax Administration). OECD Publishing.

51   OECD. (2021). BI Pre-registration form. Observatory of Public Sector Innovation. https://oecd-opsi.org/bi-pre-registration-form/

61   OECD. (2021). BI Projects Archive. https://oecd-opsi.org/bi-projects/

19   OECD. (2021). Enhancing economic performance and well-being in Chile: Policy Actions for affordable and accessible pharmaceuticals. OECD Publishing. https://www.oecd.org/economy/surveys/CHL_OECD_policy_actions_affordable_and_accessible_pharmaceuticals.pdf

67   OECD. (2021). Evidence-based and data-driven public communication. In OECD Report on Public Communication The Global Context and the Way Forward. OECD Publishing. https://www.oecd-ilibrary.org/sites/22f8031c-en/1/1/index.html?itemId=/content/publication/22f8031c-en&_csp_=010df73ac7c672500399f38b883dde4f&itemIGO=oecd&itemContentType=book

12   OECD. (2021). Good Practice Principles for Data Ethics in the Public Sector. OECD Publishing. https://www.oecd.org/digital/digital-government/good-practice-principles-for-data-ethics-in-the-public-sector.htm

11   OECD. (2021). Recommendation of the Council for Agile Regulatory Governance to Harness Innovation. OECD Legal Instruments; OECD/LEGAL/0464. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0464

31   OECD. (2021). Roundtable on Dark Commercial Patterns Online—Summary of Discussion. Roundtable on Dark Commercial Patterns Online, Paris. https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/CP(2020)23/FINAL&docLanguage=En

53   Office of Evaluation Sciences. (n.d.). Evaluations. Retrieved April 11, 2022, from https://oes.gsa.gov/work/

60   Office of Evaluation Sciences. (n.d.). OES Project Process. Retrieved April 11, 2022, from https://oes.gsa.gov/projectprocess/

59   Office of Evaluation Sciences. (n.d.). Unexpected and Null Results Can Help Build Federal Evaluation Plans and Learning Agendas. Unexpected and Null Results Can Help Build Federal Evaluation Plans and Learning Agendas. https://oes.gsa.gov/assets/files/unexpected-results-2-pager.pdf

18   Office of Evaluation Sciences. (2021). A Behavioral Insights Guide for Improving Payment Integrity. US Chief Financial Officers Council. https://oes.gsa.gov/collaborations/improper-payments-playbook/

37   Ontario Ministry of Health. (2016). Behavioural Insights Pilot Project – Organ Donor Registration. Ontario.Ca. http://www.ontario.ca/page/behavioural-insights-pilot-project-organ-donor-registration

49   Pearson, N., Naylor, P.-J., Ashe, M. C., Fernandez, M., Yoong, S. L., & Wolfenden, L. (2020). Guidance for conducting feasibility and pilot studies for implementation trials. Pilot and Feasibility Studies, 6(1), 167. https://doi.org/10.1186/s40814-020-00634-w

66   Saeri, A. K., Slattery, P., Tear, M. J., Varazzani, C., Epstein, D., Knott, C., Kusmanoff, A., Knott, C., Bagshaw, H., Phillips, K., & Liao, J. (2021). Scale up of behaviour change interventions: A rapid review of evidence and practice. OSF Preprints. https://doi.org/10.31219/osf.io/scd3k

68   Service, O., Hallsworth, M., Halpern, D., Algate, F., Gallagher, R., Nguyen, S., Ruda, S., & Sandards, M. (2012). EAST: Four Simple Ways to Apply Behavioural Insights. The Behavioural Insights Team. https://www.bi.team/publications/east-four-simple-ways-to-apply-behavioural-insights/

64   Shah, A., Osborne, M., Lefkowitz, J., Fertig, A., Soman, D., & Mazar, N. (2018). Can Making Family Salient Improve Retirement Contributions? Evidence From Field Experiments in Mexico. ACR North American Advances, NA-46. https://www.acrwebsite.org/volumes/2411999/volumes/v46/NA-46

4   Sunstein, C. (2015). The ethics of nudging. Yale J. on Reg., 32, p.413.27. Sunstein, C. (2018). Sludge and Ordeals. Duke Law Journal, 68, 1843. https://doi.org/

28   Sunstein, C. (2020). Sludge Audits. Behavioural Public Policy, 1–20. https://doi.org/10.1017/bpp.2019.32

44   Sunstein, C. (2021). Sludge: What Stops Us from Getting Things Done. MIT Press.

29   Thaler, R. (2018). Nudge, not sludge. Science, 361(6401), 431.

6   Thaler, R., & Sunstein, C. (2008). Nudge: Improving decisions about health, wealth, and happiness. (first). Yale University Press.

41   The American Psychological Association. (2003). Ethical principles of psychologists and code of conduct. The American Psychological Association. https://www.apa.org/ethics/code

50   UK Government Communication Service. (2021). IN CASE: A behavioural approach to anticipating unintended consequences. Government of United Kingdom. https://gcs.civilservice.gov.uk/publications/in-case-a-behavioural-approach-to-anticipating-unintended-consequences/

40   Varga, P. (2018). OGP's Participation and Co-Creation Toolkit: From usual suspects to business as usual. Open Government Partnership. https://www.opengovpartnership.org/stories/ogps-participation-and-co-creation-toolkit-from-usual-suspects-to-business-as-usual/

69   Wein, T., Schilling, M., Hammond, P., Mumo, J., & Juma, C. (2021). Value and validation. How feedback enhances the quality of research outputs. The Busara Blog. https://medium.com/busara-center-blog/value-and-validation-113750e7c0ad

22   World Bank Group. (2015). World Development Report 2015: Mind, Society, and Behavior. World Bank. https://www.worldbank.org/en/publication/wdr2015

38   Zion & Zion. (n.d.). Campaign Drives Record Breaking Registrations. Campaign Drives Record Breaking Registrations. Retrieved April 11, 2022, from https://www.zionandzion.com/case-study/campaign-drives-record-breaking-registrations/

OECD ⟫⟫ OECD

OPSI | Observatory of
      | Public Sector
      | Innovation